

Db 649 FPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRV 708
 QY 102 VSVLTVLHQDLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQ 161
 Db 709 VSVLTVLHQDLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQ 768
 QY 162 VSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNV 221
 Db 769 VSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNV 828
 QY 222 FSCSVMEALHNHYTQKSLSLSPCK 246
 Db 829 FSCSVMEALHNHYTQKSLSLSPCK 853

RESULT 14
 AAY92184
 ID AAY92184 standard; protein; 859 AA.
 XX
 AC AAY92184;
 XX
 DT 01-AUG-2000 (first entry)
 XX
 DE Human gpl30-Fc-His6.
 XX
 KW gpl30-Fc-His6; cytokine; antagonist; CNTF; receptor; fusion protein;
 KW cytostatic; immunomodulator; osteopathic.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX

Key Location/Qualifiers
 FH Protein 1..619
 FT /label= gp130
 FT Peptide 1..22
 FT /label= signal_peptide
 FT Peptide 620..621
 FT /note= "Ser-Gly bridge"
 FT Disulfide-bond 632
 FT /note= "forms inter-chain disulfide bridge that link two FC domains"
 FT Disulfide-bond 635
 FT /note= "forms inter-chain disulfide bridge that link two FC domains"
 FT Protein 662..853
 FT /label= IgG1_Fc_domain
 FT Peptide 854..859
 FT /label= histidine_tag

XX WO200018932-A2.
 XX
 XX
 XX 06-APR-2000.
 XX
 XX 22-SEP-1999; 99WO-US22045.
 XX
 XX 25-SEP-1998; 98US-0101858.
 XX 19-MAY-1999; 99US-0313942.
 XX
 XX (REGE-) REGENERON PHARM INC.
 XX
 XX Stahl N, Yancopoulos GD;
 XX
 XX WPI; 2000-293165/25.
 XX
 XX Isolated nucleic acid molecule for treating cytokine-related diseases
 PT or disorders encodes a fusion polypeptide capable of binding a cytokine
 PT to form a nonfunctional complex
 XX
 XX Example 3; Fig 4; 152pp; English.
 XX
 XX The invention concerns production of antagonists to any cytokine that
 CC utilizes an alpha specificity determining component, which when combined
 CC with the cytokine, binds to a first beta signal transducing component to

CC form a non-functional intermediate which then binds to a second beta
 CC signal transducing component causing beta-receptor dimerization, the
 CC soluble alpha specificity determining component of the receptor
 CC (SR-alpha) and the extracellular domain of the first beta signal
 CC transducing component of the cytokine receptor (beta-1) are combined to
 CC form heterodimers (SR-alpha:beta-1) that act as antagonist to the
 CC cytokine by binding the cytokine to form a non-functional complex. The
 CC receptor components are shared by cytokines such as the CNTF (ciliary
 CC neurotrophic factor) family of cytokines. The invention provides the
 CC basis for the development of IL-6 antagonists, as they show that if, in
 CC the presence of a ligand, a non-functional intermediate complex,
 CC consisting of the ligand, its alpha receptor and its beta-1 receptor
 CC component, can be formed, it will effectively block the action of the
 CC ligand. Effective antagonists of IL-6 or CNTF consist of heterodimers
 CC of the extracellular domains of the alpha specificity determining
 CC components of their receptors and the extracellular domain of gp130.
 CC The resultant heterodimers, function as high-affinity traps, rendering
 CC the cytokine inaccessible to form a signal transducing complex with the
 CC native membrane-bound forms of their receptor. The nucleic acids and
 CC polypeptides are useful for treating cytokine-related diseases or
 CC disorders such as osteoporosis and primary and secondary effects of
 CC cancer including multiple myeloma or cachexia.
 XX

SQ Sequence 859 AA;
 Query Match 94.0%; Score 1254; DB 21; Length 859;
 Best Local Similarity 60.5%; Pred. No. 1e-35;
 Matches 233; Conservative 8; Mismatches 5; Indels 139; Gaps 3;

QY 1 DW-----LKAF----- 6
 ||
 Db 469 DWQEDGTVHRTVLRGNLAESKCYLITVPVYADGPGSPESIKAYLKQAPPKGTVRTK 528
 QY 7 -----YDKVAEKLKEAFM----- 19
 : : : :
 Db 529 KVGKNEAVLEWDQLPVDVQNGFIRNYTIFYRTIIGNETAVNVDSHTEYTLSSLTSDTL 588
 QY 20 -----DKTHTCPCPAPELLGSPSVFL 41
 |||||
 Db 589 MVRMAAYTDGGKDGPEFTTTPKFAQGEIESGEPKSCDKTHCPCPAPELLGSPSVFL 648
 QY 42 FPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRV 101
 |||||
 Db 649 FPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRV 708
 QY 102 VSVLTVLHQDLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQ 161
 |||||
 Db 709 VSVLTVLHQDLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQ 768
 QY 162 VSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNV 221
 |||||
 Db 769 VSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNV 828
 QY 222 FSCSVMEALHNHYTQKSLSLSPCK 246
 |||||
 Db 829 FSCSVMEALHNHYTQKSLSLSPCK 853

RESULT 15
 AAY92205
 ID AAY92205 standard; Protein; 1158 AA.
 XX
 AC AAY92205;
 XX
 DT 01-AUG-2000 (first entry)
 XX
 DE Fusion polypeptide 616, IL-6 trap.
 XX
 KW IL-6 trap; cytokine; antagonist; CNTF; receptor; fusion protein;
 KW cytostatic; immunomodulator; osteopathic.
 XX
 OS Synthetic.
 OS Homo sapiens.

XX PN WO200018932-A2.
XX PD 06-APR-2000.
XX PF 22-SEP-1999; 99WO-US22045.
XX PR 25-SEP-1998; 98US-0101858.
XX PR 19-MAY-1999; 98US-0313942.
XX PA (REGE-) REGENERON PHARM INC.
XX PI Stahl N, Yancopoulos GD;
XX DR WPI; 2000-293165/25.
XX DR N-PSDB; AAA09047.
XX PT Isolated nucleic acid molecule for treating cytokine-related diseases
XX PT or disorders encodes a fusion polypeptide capable of binding a cytokine
XX PT to form a nonfunctional complex
XX PS Example 6; Fig 25A-F; 152pp; English.
XX CC This sequence shows fusion polypeptide 616, which is capable of
XX CC binding cytokine IL-6 to form a non-functional complex.
XX CC The invention concerns production of antagonists to any cytokine that
XX CC utilizes an alpha specificity determining component, which when combined
XX CC with the cytokine, binds to a first beta signal transducing component to
XX CC form a non-functional intermediate which then binds to a second beta
XX CC signal transducing component causing beta-receptor dimerization, the
XX CC soluble alpha specificity determining component of the receptor
XX CC (sr-alpha) and the extracellular domain of the first beta signal
XX CC transducing component of the cytokine receptor (beta-1) are combined to
XX CC form heterodimers (sr-alpha:beta-1) that act as antagonist to the
XX CC cytokine by binding the cytokine to form a non-functional complex. The
XX CC receptor components are shared by cytokines such as the CNF (ciliary
XX CC neurotrophic factor) family of cytokines. The invention provides the
XX CC basis for the development of IL-6 antagonists, as they show that if, in
XX CC the presence of a ligand, a non-functional intermediate complex,
XX CC consisting of the ligand, its alpha receptor and its beta-1 receptor
XX CC component, can be formed, it will effectively block the action of the
XX CC ligand. Effective antagonists of IL-6 or CNF consist of heterodimers
XX CC of the extracellular domains of the alpha specificity determining
XX CC components of their receptors and the extracellular domain of gp130.
XX CC The resultant heterodimers, function as high-affinity traps, rendering
XX CC the cytokine inaccessible to form a signal transducing complex with the
XX CC native membrane-bound forms of their receptor. The nucleic acids and
XX CC polypeptides are useful for treating cytokine-related diseases or
XX CC disorders such as osteoporosis and primary and secondary effects of
XX CC cancer including multiple myeloma or cachexia.
XX SQ Sequence 1158 AA;

Query Match 94.0%; Score 1254; DB 21; Length 1158;
Best Local Similarity 61.2%; Pred. No. 2e-35;
Matches 233; Conservative 8; Mismatches 5; Indels 135; Gaps 3;
QY 1 DW-----LKA-----6
DB 778 DWQDQGVHRTVLRNLAKSCYLITVTPVYADGPGSPESIKAYLKQAPPSKGTVRTK 837
QY 7 -----YDKVAEKLKAF-----19
DB 838 KVGKNEAVLEWOLPVDVQNGFIRNVTIFYRIIGNETAVNDSSHTEYTLSSLTSDTL 897
QY 20 -----DKTHTCPCPAPPELLGGPSVFLFPPK 45
DB 898 MVRMAAYTDEGGKDGPEFTTTPKFAQGEIESGGDKTHTCPCPAPPELLGGPSVFLFPPK 957
QY 46 PKDTLMISRTPEVTCVVVDVSHEDPEVKFNKYVDGVVHNNAKTKPREQYNSTYRVSVL 105
DB 958 PKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVVHNNAKTKPREQYNSTYRVSVL 1017

QY 106 TVLHODWLNKGEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNOVSLT 165
DB 1018 TVLHODWLNKGEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNOVSLT 1077
QY 166 CLVKGFPYSDIAVEWESNGQPENNYKTTTPVLDSDGSEFFLYSKLTVDKSRWQOGNVSFCS 225
DB 1078 CLVKGFPYSDIAVEWESNGQPENNYKTTTPVLDSDGSEFFLYSKLTVDKSRWQOGNVSFCS 1137
QY 226 VVHEALHNHYTQKSLSLSPGK 246
DB 1138 VVHEALHNHYTQKSLSLSPGK 1158

Search completed: April 21, 2003, 10:45:22
Job time : 32.9412 secs

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:26 ; Search time 13.7471 seconds
(without alignments)
526.515 Million cell updates/sec

Title: 7LINK2

Perfect score: 1334

Sequence: 1 DWLKAFYDKVAEKLKEAFMD.....MHEALHNYTKSLSPGK 246

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1258	94.3	476	3	US-08-487-550-12
2	1252	93.9	449	1	US-08-458-516-13
3	1250	93.7	476	2	US-08-378-939-10
4	1249	93.6	347	1	US-07-940-861-43
5	1249	93.6	347	1	US-08-459-512-43
6	1249	93.6	347	2	US-08-459-657-43
7	1249	93.6	347	2	US-08-460-132-43
8	1249	93.6	347	4	US-08-466-465-8
9	1249	93.6	347	5	PCT-US92-02050-43
10	1249	93.6	452	3	US-09-027-449-71
11	1249	93.6	452	4	US-09-026-985-71
12	1249	93.6	452	4	US-09-121-952A-71
13	1249	93.6	452	4	US-09-234-340A-71
14	1249	93.6	459	1	US-08-157-101A-7
15	1249	93.6	664	3	US-08-957-063-16
16	1249	93.6	664	3	US-08-957-063-18
17	1249	93.6	664	4	US-09-487-685-16
18	1249	93.6	664	4	US-09-487-685-18
19	1249	93.6	664	4	US-08-802-805D-16
20	1249	93.6	664	4	US-08-802-805D-18
21	1248	93.6	476	3	US-08-487-550-4
22	1247	93.5	437	5	PCT-US96-10043-11
23	1247	93.5	446	3	US-08-397-411-7
24	1246	93.4	454	2	US-07-934-373C-22
25	1246	93.4	454	3	US-08-437-642B-22
26	1246	93.4	454	4	US-08-146-206C-22
27	1246	93.4	454	5	PCT-US93-07832-22

28	1245	93.3	371	1	US-08-236-311-7	Sequence 7, Appli
29	1245	93.3	371	3	US-08-457-918-7	Sequence 7, Appli
30	1245	93.3	388	4	US-09-131-247-16	Sequence 16, Appl
31	1245	93.3	424	5	PCT-US93-03866-12	Sequence 12, Appl
32	1245	93.3	424	5	PCT-US93-03866-14	Sequence 14, Appl
33	1245	93.3	451	2	US-08-887-352B-14	Sequence 14, Appl
34	1245	93.3	451	2	US-08-887-352B-16	Sequence 16, Appl
35	1245	93.3	451	2	US-08-887-352B-18	Sequence 18, Appl
36	1245	93.3	451	3	US-08-466-151-65	Sequence 65, Appl
37	1245	93.3	451	4	US-09-109-207C-14	Sequence 14, Appl
38	1245	93.3	451	4	US-09-109-207C-16	Sequence 16, Appl
39	1245	93.3	451	4	US-09-109-207C-18	Sequence 18, Appl
40	1245	93.3	451	4	US-09-282-505-2	Sequence 2, Appli
41	1245	93.3	451	4	US-09-054-255-2	Sequence 14, Appl
42	1245	93.3	451	4	US-09-296-005-14	Sequence 16, Appl
43	1245	93.3	451	4	US-09-296-005-16	Sequence 16, Appl
44	1245	93.3	451	4	US-09-296-005-18	Sequence 18, Appl
45	1245	93.3	453	3	US-08-466-151-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-08-487-550-12
; Sequence 12, Application US/08487550
; Patent No. 6113898
; GENERAL INFORMATION:
; APPLICANT: Anderson, Darrell R.
; TITLE OF INVENTION: "MONKEY MONOCLONAL ANTIBODIES SPECIFIC
; TITLE OF INVENTION: TO HUMAN B7.1 AND/OR B7.2 PRIMATIZED FORMS THEREOF,
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING, AND USE THEREOF AS
; TITLE OF INVENTION: IMMUNOSUPPRESSANTS"
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: 699 Prince Street
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22314

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/487,550
APPLICATION NUMBER: US/08/487,550
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 012712-131
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6620
TELEFAX: 703-836-2021
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 476 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-487-550-12

Query Match 94.3%; Score 1258; DB 3; Length 476;
Best Local Similarity 56.7%; Pred. No. 3.4e-44;
Matches 233; Conservative 5; Mismatches 8; Indels 165; Gaps 1;

QY 1 DWLKAFYD-----
:|: :||

Db 66 EWIGSFYSSGNTYNPFLSKSVTISTDTSKNQFLKLSMTAADTAVYCVDRDLFSW 125

```
QY 9 ----- 8
Db 126 GMVNNWFDWGPGLVLTSSASTKGPSVFPLAPSKSTSGGTAALGCLVKDYFPEPTV 185
QY 9 -----KVAEKLK 15
Db 186 SWNSGALTSVHTFPVAVLOSSGLYSLSVVTVPSSSLGTQTYICNVNHNKPSNTKVDKXAE 245
QY 16 EAFMDKTHTCPCPCAPPELLGGPSVFLFPPKPKDTHLMSRTPEVTCVVVDVSHEDPEVKEN 75
Db 246 PKSCDKTHTCPCPCAPPELLGGPSVFLFPPKPKDTHLMSRTPEVTCVVVDVSHEDPEVKEN 305
QY 76 WYVDGVEVHNKTKPREEQYNSTYRVSVTLVHODWLNKGEYCKCKVSNKALPAPIEKTI 135
Db 306 WYVDGVEVHNKTKPREEQYNSTYRVSVTLVHODWLNKGEYCKCKVSNKALPAPIEKTI 365
QY 136 SKAGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYPSPDIAVEWESNGQPENNYKTTPP 195
Db 366 SKAGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYPSPDIAVEWESNGQPENNYKTTPP 425
QY 196 VLDSGDSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTOKLSLSLSPGK 246
Db 426 VLDSGDSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTOKLSLSLSPGK 476

RESULT 2
US-08-458-516-13
; Sequence 13, Application US/08458516
; Patent No. 5777085
; GENERAL INFORMATION:
; APPLICANT: Co, Man Sung
; APPLICANT: Tso, J. Yun
; TITLE OF INVENTION: Humanized Antibodies Reactive with
; TITLE OF INVENTION: GPIIB/IIIA
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: William M. Smith
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,516
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION NUMBER: US 08/059,159
; APPLICATION NUMBER: US 08/059,159
; FILING DATE: 03-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11823-37-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 449 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-458-516-13

Query Match 93.9%; Score 1252; DB 1; Length 449;
Best Local Similarity 57.4%; Pred. No. 5.3e-44;
Matches 232; Conservative 5; Mismatches 9; Indels 158; Gaps 1;
```

```
QY 1 DMLKAFY ----- 7
Db 46 EMIGVLYPGSGGTNYNEKEFKGRVTLTVDESTNTAYMELSSLSRSEDPAVYFCARRDGNYGW 105
QY 8 ----- 7
Db 106 FAYWGGTGLTVTSSASTKGPSVFPLAPSKSTSGGTAALGCLVKDYFPEPTVTVSMNSGAL 165
QY 8 -----DKVAEKLKEAFMDKT 22
Db 166 TSGVHTFPVAVLOSSGLYSLSVVTVPSSSLGTQTYICNVNHNKPSNTKVDKKEPKSCDKT 225
QY 23 HFCPPCPAPPELLGGPSVFLFPPKPKDTHLMSRTPEVTCVVVDVSHEDPEVKFNWYVDGVE 82
Db 226 HFCPPCPAPPELLGGPSVFLFPPKPKDTHLMSRTPEVTCVVVDVSHEDPEVKFNWYVDGVE 285
QY 83 VHNKTKPREEQYNSTYRVSVTLVHODWLNKGEYCKCKVSNKALPAPIEKTIISKAGQP 142
Db 286 VHNKTKPREEQYNSTYRVSVTLVHODWLNKGEYCKCKVSNKALPAPIEKTIISKAGQP 345
QY 143 REPOVYTLPPSRDELTKNOVSLTCLVKGYPSPDIAVEWESNGQPENNYKTTPPVLDSDGS 202
Db 346 REPOVYTLPPSRDELTKNOVSLTCLVKGYPSPDIAVEWESNGQPENNYKTTPPVLDSDGS 405
QY 203 FFYLSKLTVDKSRWQGNVFCSCVMHEALHNHYTOKLSLSLSPGK 246
Db 406 FFYLSKLTVDKSRWQGNVFCSCVMHEALHNHYTOKLSLSLSPGK 449

RESULT 3
US-08-378-939-10
; Sequence 10, Application US/08378939
; Patent No. 5876961
; GENERAL INFORMATION:
; APPLICANT: CROWE, JAMES SCOTT
; APPLICANT: LEWIS, ALAN PETER
; TITLE OF INVENTION: PRODUCTION OF ANTIBODIES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ
; STREET: 555 THIRTEENTH ST. N.W.
; CITY: WASHINGTON
; STATE: D. C.
; COUNTRY: U. S.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/378,939
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/952640
; FILING DATE: 01-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: ERNST, BARBARA G
; REGISTRATION NUMBER: 30,377
; REFERENCE/DOCKET NUMBER: 1808-118
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 783-6040
; TELEFAX: (202) 783-6031
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 476 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-378-939-10
```

```
Query Match 93.7%; Score 1250; DB 2; Length 476;
Best Local Similarity 56.6%; Pred. No. 7.4e-44;
Matches 233; Conservative 7; Mismatches 6; Indels 166; Gaps 2;

QY 1 DWL-----KAF 6
Db 65 ENWGIIPLFGTPTYSQFGRVITADKSTSTAHMELTSLRSEDPAVYVCATDRYQAN 124
QY 7 YD-----8
Db 125 FDRARVGFDPWCGTLTVVSSASTKGPSVFPPLAPSSKSTSGTAALGCLVKDYFPEPVT 184
QY 9 -----KVAEKL 14
Db 185 VSNNGALTSVGHFFPAVLQSSGLYSLSSVTVTPSSSLGTQTQYICNVNHNKPSNTKDKV 244
QY 15 KEAFMDKTHTCPCPAPELLGGPSVFLPPKPKDMLMISRTPEVTCVVVDVSHEDPEVK 74
Db 245 EPKSCDKTHTCPCPAPELLGGPSVFLPPKPKDMLMISRTPEVTCVVVDVSHEDPEVK 304
QY 75 NWYVDGVEVHNATKPREEQYNSTYRVSVLTVLHODWLNKGYCKVSNKALPAPIET 134
Db 305 NWYVDGVEVHNATKPREEQYNSTYRVSVLTVLHODWLNKGYCKVSNKALPAPIET 364
QY 135 ISKAKQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTP 194
Db 365 ISKAKQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTP 424
QY 195 PVLDSDGSFFLYSKLTVDKSRWQGNVFSCSVMEALHNHYTKLSLSPGK 246
Db 425 PVLDSDGSFFLYSKLTVDKSRWQGNVFSCSVMEALHNHYTKLSLSPGK 476

RESULT 4
US-07-940-861-43
; Sequence 43, Application US/07940861
; Patent No. 5547853
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: WALLNER, Barbara P.
; APPLICANT: MILLER, Glenn T.
; APPLICANT: ROSA, Margaret D.
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 875 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022-6250
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/940,861
; FILING DATE: 21-OCT-1992
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02050
; FILING DATE: 12-MAR-1992
; APPLICATION DATA:
; FILING DATE: 07/667,971
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,967
; FILING DATE: 07-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY, James F., Jr.
; REGISTRATION NUMBER: 27,794
```

```
REFERENCE/DOCKET NUMBER: B151CIP2
TELEPHONE: (212)715-0600
TELEFAX: (212)715-0673
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 347 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-940-861-43

Query Match 93.6%; Score 1249; DB 1; Length 347;
Best Local Similarity 80.2%; Pred. No. 3.7e-44;
Matches 235; Conservative 1; Mismatches 8; Indels 49; Gaps 2;

QY 2 WLKAFDYKVAEKLKEAF-----18
Db 56 WKQK-KDKVAELENSEFRASFKNRVYLDTVSGSLTIYNLTSSDEDEYEMESPNITDTM 114
QY 19 -----MDKTHTCPCPAPELLGGPSVFLPPKPKDMLMISRTPEVTCVVVDVSHEDPEVK 73
Db 115 KFFLYVDKTHTCPCPAPELLGGPSVFLPPKPKDMLMISRTPEVTCVVVDVSHEDPEVK 174
QY 74 FMYVDGVEVHNATKPREEQYNSTYRVSVLTVLHODWLNKGYCKVSNKALPAPIEK 133
Db 175 FMYVDGVEVHNATKPREEQYNSTYRVSVLTVLHODWLNKGYCKVSNKALPAPIEK 234
QY 134 TISKAKQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 193
Db 235 TISKAKQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 294
QY 194 PVLDSDGSFFLYSKLTVDKSRWQGNVFSCSVMEALHNHYTKLSLSPGK 246
Db 295 PVLDSDGSFFLYSKLTVDKSRWQGNVFSCSVMEALHNHYTKLSLSPGK 347

RESULT 5
US-08-459-512-43
; Sequence 43, Application US/08459512
; Patent No. 5728677
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: WALLNER, Barbara P.
; APPLICANT: MILLER, Glenn T.
; APPLICANT: ROSA, Margaret D.
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 875 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022-6250
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,512
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02050
; FILING DATE: 12-MAR-1992
; APPLICATION NUMBER: US 07/667,971
; FILING DATE: 12-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,967
```

;; FILING DATE: 07-OCT-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: HALEY, James F., Jr.
;; REGISTRATION NUMBER: 27,794
;; REFERENCE/DOCKET NUMBER: B151CIP2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212)715-0600
;; TELEFAX: (212)715-0673
;; TELEX: 14-8367
;; INFORMATION FOR SEQ ID NO: 43:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 347 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-459-512-43

Query Match 93.6%; Score 1249; DB 1; Length 347;

Best Local Similarity 80.2%; Pred. No. 3.7e-44;

Matches 235; Conservative 1; Mismatches 8; Indels 49; Gaps 2;

QY 2 WLKAFYDKVAEKLKEAF----- 18
DB 56 WKQK-KDKVAELENSEFRASFKNRVYDVTVSGSLTIYNTLSDEDEYEMESPNITDTM 114
QY 19 -----MDKTHTCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 73
DB 115 KFFLYVDKTHTCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 174
QY 74 FMYVDGVEVHNATKPREEQYNSTYRVSVLTVLHQDLNKGKEYCKVSNKALPAIEK 133
DB 175 FMYVDGVEVHNATKPREEQYNSTYRVSVLTVLHQDLNKGKEYCKVSNKALPAIEK 234
QY 134 TISKAGGPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVEMESNGOPENNYKTT 193
DB 235 TISKAGGPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVEMESNGOPENNYKTT 294
QY 194 PPVLDSGDSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPGK 246
DB 295 PPVLDSGDSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPGK 347

RESULT 6

US-08-459-657-43
; Sequence 43, Application US/08459657
; Patent No. 5914111

GENERAL INFORMATION:

;; APPLICANT: BIOGEN, INC.
;; APPLICANT: WALLNER, Barbara P.
;; APPLICANT: MILLER, Glenn T.
;; APPLICANT: ROSA, Margaret D.
;; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
;; TITLE OF INVENTION: FUNCTION ASSOCIATED ANTIGEN 3
;; NUMBER OF SEQUENCES: 43
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Fish & Neave
;; STREET: 875 Third Avenue
;; CITY: New York
;; STATE: New York
;; COUNTRY: U.S.A.
;; ZIP: 10022-6250

COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/459,657
;; FILING DATE: 02-JUN-1995
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US92/02050
;; FILING DATE: 12-MAR-1992

;; APPLICATION NUMBER: US 07/667,971
;; FILING DATE: 12-NAR-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/770,967
;; FILING DATE: 07-OCT-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: HALEY, James F., Jr.
;; REGISTRATION NUMBER: 27,794
;; REFERENCE/DOCKET NUMBER: B151CIP2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212)715-0600
;; TELEFAX: (212)715-0673
;; TELEX: 14-8367
;; INFORMATION FOR SEQ ID NO: 43:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 347 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-459-657-43

Query Match 93.6%; Score 1249; DB 2; Length 347;

Best Local Similarity 80.2%; Pred. No. 3.7e-44;

Matches 235; Conservative 1; Mismatches 8; Indels 49; Gaps 2;

QY 2 WLKAFYDKVAEKLKEAF----- 18
DB 56 WKQK-KDKVAELENSEFRASFKNRVYDVTVSGSLTIYNTLSDEDEYEMESPNITDTM 114
QY 19 -----MDKTHTCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 73
DB 115 KFFLYVDKTHTCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 174
QY 74 FMYVDGVEVHNATKPREEQYNSTYRVSVLTVLHQDLNKGKEYCKVSNKALPAIEK 133
DB 175 FMYVDGVEVHNATKPREEQYNSTYRVSVLTVLHQDLNKGKEYCKVSNKALPAIEK 234
QY 134 TISKAGGPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVEMESNGOPENNYKTT 193
DB 235 TISKAGGPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVEMESNGOPENNYKTT 294
QY 194 PPVLDSGDSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPGK 246
DB 295 PPVLDSGDSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPGK 347

RESULT 7

US-08-460-132-43
; Sequence 43, Application US/08460132
; Patent No. 5928643

GENERAL INFORMATION:

;; APPLICANT: BIOGEN, INC.
;; APPLICANT: WALLNER, Barbara P.
;; APPLICANT: MILLER, Glenn T.
;; APPLICANT: ROSA, Margaret D.
;; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
;; TITLE OF INVENTION: FUNCTION ASSOCIATED ANTIGEN 3
;; NUMBER OF SEQUENCES: 43
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Fish & Neave
;; STREET: 875 Third Avenue
;; CITY: New York
;; STATE: New York
;; COUNTRY: U.S.A.
;; ZIP: 10022-6250

COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/460,132
;; FILING DATE: 02-JUN-1995

```

; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/940,861
; FILING DATE: 21-OCT-1992
; APPLICATION NUMBER: PCT/US92/02050
; FILING DATE: 12-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/667,971
; FILING DATE: 12-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,967
; FILING DATE: 07-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY, James F., Jr.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B151CIP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)715-0600
; TELEFAX: (212)715-0673
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-460-132-43

Query Match 93.6%; Score 1249; DB 2; Length 347;
Best Local Similarity 80.2%; Pred. No. 3.7e-44;
Matches 235; Conservative 1; Mismatches 8; Indels 49; Gaps 2;

QY 2 WLKAFYDKVAEKLKEAF-----18
DB 56 WKQK-KDKVAELENSEFRASFKNRYLDTVSGSLTYNLTSSDEDEYEMESPNTDTM 114
QY 19 -----MDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVK 73
DB 115 KFFLYVDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVK 174
QY 74 FNNYVDGVEVHNKTKPREEQYNSTYRVSVLTVLHODWLNKREYKCKVSNKALPAPIEK 133
DB 175 FNNYVDGVEVHNKTKPREEQYNSTYRVSVLTVLHODWLNKREYKCKVSNKALPAPIEK 234
QY 134 TISKAKQPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 193
DB 235 TISKAKQPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 294
QY 194 PPVLDSGSEFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGK 246
DB 295 PPVLDSGSEFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGK 347

RESULT 8
US-08-466-465-8
; Sequence 8, Application US/08466465
; Patent No. 6162432
; GENERAL INFORMATION:
; APPLICANT: Wallner, Barbara P.
; APPLICANT: Cooper, Kevin D.
; TITLE OF INVENTION: Method of Prophylaxis or Treatment of Antigen
; TITLE OF INVENTION: Presenting Cell Driven Skin Conditions Using
; TITLE OF INVENTION: Inhibitors of the CD2/LFA-3 Interaction
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,465
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08755
; FILING DATE: 06-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/862,022
; FILING DATE: 12-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,969
; FILING DATE: 07-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis (PLM)
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGP-111CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-466-465-8

Query Match 93.6%; Score 1249; DB 4; Length 347;
Best Local Similarity 80.2%; Pred. No. 3.7e-44;
Matches 235; Conservative 1; Mismatches 8; Indels 49; Gaps 2;

QY 2 WLKAFYDKVAEKLKEAF-----18
DB 56 WKQK-KDKVAELENSEFRASFKNRYLDTVSGSLTYNLTSSDEDEYEMESPNTDTM 114
QY 19 -----MDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVK 73
DB 115 KFFLYVDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVK 174
QY 74 FNNYVDGVEVHNKTKPREEQYNSTYRVSVLTVLHODWLNKREYKCKVSNKALPAPIEK 133
DB 175 FNNYVDGVEVHNKTKPREEQYNSTYRVSVLTVLHODWLNKREYKCKVSNKALPAPIEK 234
QY 134 TISKAKQPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 193
DB 235 TISKAKQPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 294
QY 194 PPVLDSGSEFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGK 246
DB 295 PPVLDSGSEFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGK 347

RESULT 9
PCT-US92-02050-43
; Sequence 43, Application PC/TUS9202050
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: WALLNER, Barbara P.
; APPLICANT: MILLER, Glenn T.
; APPLICANT: ROSA, Margaret D.
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
; TITLE OF INVENTION: FUNCTION ASSOCIATED ANTIGEN 3
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 875 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022-6250

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/02050
FILING DATE: 19920312
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/667,971
FILING DATE: 12-MAR-1991
APPLICATION DATA:
FILING DATE: 07-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: HALEY, James F., Jr.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B151CIP2
TELEPHONE: (212)715-0600
TELEFAX: (212)715-0673
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 347 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-02050-43

Query Match 93.6%; Score 1249; DB 5; Length 347;
Best Local Similarity 80.2%; Pred. No. 3.7e-44;
Matches 235; Conservative 1; Mismatches 8; Indels 49; Gaps 2;

QY 2 WLKAFYDKVAEKLKEAF-----18
DB 56 WKQK-KDKVAELENSEFRAFSSFKNRVYLDVTVSGSLTIYNLTSSDEDEYEMESPNTIDPM 114
QY 19 -----MDKTHTCPPCPAPPELLGGPSVFLFPPPKPKDTLMISRTPEVTCVVDVSHEDPEVK 73
DB 115 KFFLYVDKTHTCPPCPAPPELLGGPSVFLFPPPKPKDTLMISRTPEVTCVVDVSHEDPEVK 174
QY 74 FNNYDGVGVHNAKTKPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 133
DB 175 FNNYDGVGVHNAKTKPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 234
QY 134 TISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKT 193
DB 235 TISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKT 294
QY 194 PPVLDSGDSFFLYSKLTVDKSRWQOGNVFSCSVMEALHNHYTQKSLSLSPGK 246
DB 295 PPVLDSGDSFFLYSKLTVDKSRWQOGNVFSCSVMEALHNHYTQKSLSLSPGK 347

RESULT 10
US-09-027-449-71
Sequence 71, Application US/09027449
Patent No. 6025158
GENERAL INFORMATION:
APPLICANT: Gonzalez, Tania R.
APPLICANT: Leong, Steven R.
APPLICANT: Presta, Leonard G.
TITLE OF INVENTION: Antibody Fragment-Polymer Conjugates and
TITLE OF INVENTION: Humanized Anti-IL-8 Monoclonal Antibodies
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA

ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/027,449
FILING DATE: 20-Feb-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/074,330
FILING DATE: 22-Jan-1998
APPLICATION DATA:
FILING DATE: 21-Feb-1997
ATTORNEY/AGENT INFORMATION:
NAME: Love, Richard B.
REGISTRATION NUMBER: 34,659
REFERENCE/DOCKET NUMBER: P1085R3-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-5530
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 71:
SEQUENCE CHARACTERISTICS:
LENGTH: 452 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-09-027-449-71

Query Match 93.6%; Score 1249; DB 3; Length 452;
Best Local Similarity 66.3%; Pred. No. 7.2e-44;
Matches 232; Conservative 6; Mismatches 5; Indels 107; Gaps 2;

QY 1 DMLKAFYD-----8
DB 106 DW---FFDWGOGTLTVSSASTKGPSVFPAPSSKSTSGGTAALGCLVKDYFPEPTVVS 162
QY 9 -----KVAEKLKE 16
DB 163 WNSGALTSGVHTFPAVLQSSGLYSLSSVTVVPSSSLGTQTYICNVNHNHPSNTKVDKVEP 222
QY 17 AFMDKTHTCPPCPAPPELLGGPSVFLFPPPKPKDTLMISRTPEVTCVVDVSHEDPEVK 76
DB 223 RSCDKTHTCPPCPAPPELLGGPSVFLFPPPKPKDTLMISRTPEVTCVVDVSHEDPEVK 282
QY 77 YVDGVEVHNAKTKPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 136
DB 283 YVDGVEVHNAKTKPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 342
QY 137 KAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPV 196
DB 343 KAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPV 402
QY 197 LDSGDSFFLYSKLTVDKSRWQOGNVFSCSVMEALHNHYTQKSLSLSPGK 246
DB 403 LDSGDSFFLYSKLTVDKSRWQOGNVFSCSVMEALHNHYTQKSLSLSPGK 452

RESULT 11
US-09-026-985-71
Sequence 71, Application US/09026985
Patent No. 6133426
GENERAL INFORMATION:
APPLICANT: Gonzalez, Tania R.
APPLICANT: Leong, Steven R.
APPLICANT: Presta, Leonard G.
TITLE OF INVENTION: Antibody Fragment-Polymer Conjugates and
TITLE OF INVENTION: Humanized Anti-IL-8 Monoclonal Antibodies
NUMBER OF SEQUENCES: 72
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way

```

; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/026,985
; FILING DATE: 20-Feb-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P1085R3-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-5530
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 452 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-09-026-985-71

Query Match 93.6%; Score 1249; DB 4; Length 452;
Best Local Similarity 66.3%; Pred. No. 7.2e-44;
Matches 232; Conservative 6; Mismatches 5; Indels 107; Gaps 2;

Qy 1 DWLKAFYD----- 8
|||
Db 106 DW---FFDVGQGLTVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVS 162
Qy 9 -----KVAEKLKE 16
|||
Db 163 WNSGALTSVHTFPAVLQSSGLYSLSSVTVTPSSSLGTQTVICNVNHNKPSNTKVDKKVEP 222
Qy 17 AFMDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVFENW 76
Db 223 KSCDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVFENW 282
Qy 77 YVDGVEVHNATKPREQYNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTIS 136
Db 283 YVDGVEVHNATKPREQYNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTIS 342
Qy 137 KAKGQPREPQVYTLPPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPV 196
Db 343 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPV 402
Qy 197 LDSGSEFLYSLKLTVDKSRWQGNVFCVSMVHEALHNHYTOKLSLSPGK 246
Db 403 LDSGSEFLYSLKLTVDKSRWQGNVFCVSMVHEALHNHYTOKLSLSPGK 452

RESULT 12
US-09-121-952A-71
; Sequence 71, Application US/09121952A
; Patent No. 6458355
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc., Hsei, Vanessa
; APPLICANT: Koumenis, Iphigenia
; APPLICANT: Leong, Steven R.
; APPLICANT: Presta, Leonard G.
; APPLICANT: Shahrokhi, Zahra
; APPLICANT: Zapata, Gerardo A.
; TITLE OF INVENTION: METHODS OF TREATING INFLAMMATORY DISEASES
; TITLE OF INVENTION: WITH ANTI-IL-8 ANTIBODY FRAGMENT-POLYMER CONJUGATES
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way

```

```

; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/121,952A
; FILING DATE: 24-Jul-1998
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/074330
; FILING DATE: 22-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/075467
; FILING DATE: 20-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P1085R4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-5530
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 452 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-09-121-952A-71

Query Match 93.6%; Score 1249; DB 4; Length 452;
Best Local Similarity 66.3%; Pred. No. 7.2e-44;
Matches 232; Conservative 6; Mismatches 5; Indels 107; Gaps 2;

Qy 1 DWLKAFYD----- 8
|||
Db 106 DW---FFDVGQGLTVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVS 162
Qy 9 -----KVAEKLKE 16
|||
Db 163 WNSGALTSVHTFPAVLQSSGLYSLSSVTVTPSSSLGTQTVICNVNHNKPSNTKVDKKVEP 222
Qy 17 AFMDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVFENW 76
Db 223 KSCDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVFENW 282
Qy 77 YVDGVEVHNATKPREQYNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTIS 136
Db 283 YVDGVEVHNATKPREQYNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTIS 342
Qy 137 KAKGQPREPQVYTLPPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPV 196
Db 343 KAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPV 402
Qy 197 LDSGSEFLYSLKLTVDKSRWQGNVFCVSMVHEALHNHYTOKLSLSPGK 246
Db 403 LDSGSEFLYSLKLTVDKSRWQGNVFCVSMVHEALHNHYTOKLSLSPGK 452

RESULT 13
US-09-234-340A-71
; Sequence 71, Application US/09234340A
; Patent No. 6468532
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc., Hsei, Vanessa
; APPLICANT: Koumenis, Iphigenia
; APPLICANT: Leong, Steven R.
; APPLICANT: Presta, Leonard G.
; APPLICANT: Shahrokhi, Zahra
; APPLICANT: Zapata, Gerardo A.

```

;; TITLE OF INVENTION: METHODS OF TREATING INFLAMMATORY DISEASES
;; TITLE OF INVENTION: WITH ANTI-IL-8 ANTIBODY FRAGMENT-POLYMER CONJUGATES
;; NUMBER OF SEQUENCES: 72
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Genentech, Inc.
;; STREET: 1 DNA Way
;; CITY: South San Francisco
;; STATE: California
;; COUNTRY: USA
;; ZIP: 94080
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: WinPatIn (Genentech)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/234,340A
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/09/121,952
;; FILING DATE: 24-Jul-1998
;; APPLICATION NUMBER: 60/074330
;; FILING DATE: 22-JAN-1998
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 60/075467
;; FILING DATE: 20-FEB-1998
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Love, Richard B.
;; REGISTRATION NUMBER: 34,659
;; REFERENCE/DOCKET NUMBER: P1085R4
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 650/225-5530
;; TELEFAX: 650/952-9881
;; INFORMATION FOR SEQ ID NO: 71:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 452 amino acids
;; TYPE: Amino Acid
;; TOPOLOGY: Linear
US-09-234-340A-71

Query Match 93.6%; Score 1249; DB 4; Length 452;
Best Local Similarity 66.3%; Pred. No. 7.2e-44;
Matches 232; Conservative 6; Mismatches 5; Indels 107; Gaps 2;

QY 1 DWLKAFYD-----KVAEKLKE 16
|| |:
Db 106 DW---FFDVGOGTLVTSSASTKGPSVPLAPSSKSTGGTAALGCLVKDYFPEPTVVS 162
QY 9 -----KVAEKLKE 16
|| |:
Db 163 WNSGALTSVHTFPAVLQSSGLYSLSSVTVPSISLGTQTYICNVNHRKPSNTKVDKVEP 222
QY 17 AFMDKTHTCPCPAPELLGGPSVFLFPPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNW 76
Db 223 KSCDKHTHTCCPCPAPELLGGPSVFLFPPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNW 282
QY 77 YVDGVEVHNAKTPREEQNSTYRVSVLTVLHQDWLNGKEYKCKVSNKALPAPTEKTTIS 136
Db 283 YVDGVEVHNAKTPREEQNSTYRVSVLTVLHQDWLNGKEYKCKVSNKALPAPTEKTTIS 342
QY 137 KAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPV 196
Db 343 KAGQPREPQVYTLPPSRDEMTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPV 402
QY 197 LDSGGSFFLYSKLTVDKSRWQOGNVSFSCSVMEALHNNHYTKQSLSPGK 246
Db 403 LDSGGSFFLYSKLTVDKSRWQOGNVSFSCSVMEALHNNHYTKQSLSPGK 452

RESULT 14
US-08-157-101A-7
; Sequence 7, Application US/08157101A

;; Patent No. 5808032
;; GENERAL INFORMATION:
;; APPLICANT: KURIHARA, TATSUYA
;; APPLICANT: MATSUKURA, SHIGEKAZU
;; APPLICANT: TSURUOKA, NOBUO
;; APPLICANT: ARIMA, KENJI
;; APPLICANT: NISHIHARA, TATSURO
;; TITLE OF INVENTION: ANTI-HBs ANTIBODY GENES AND EXPRESSION
;; TITLE OF INVENTION: PLASMIDS THEREFOR
;; NUMBER OF SEQUENCES: 9
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: PILLSBURY, MADISON & SUTRO
;; STREET: 1100 NEW YORK AVENUE, N.W.
;; CITY: WASHINGTON
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20005
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/157,101A
;; FILING DATE: 05-APR-1994
;; CLASSIFICATION: 530
;; ATTORNEY/AGENT INFORMATION:
;; NAME: TITUS, MARLANA K
;; REGISTRATION NUMBER: 35843
;; REFERENCE/DOCKET NUMBER: 9437/204199
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-861-3711
;; TELEFAX: 202-822-0944
;; TELEX: 6714627 CUCH
;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 459 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-157-101A-7

Query Match 93.6%; Score 1249; DB 1; Length 459;
Best Local Similarity 96.6%; Pred. No. 7.4e-44;
Matches 230; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 9 KVAEKLKEAFMDKTHTCPCPAPELLGGPSVFLFPPPKPKDTLMISRTPEVTCVVDVSHSHE 68
|| |:
Db 222 KYDKKVEPKSCDKHTHTCCPCPAPELLGGPSVFLFPPPKPKDTLMISRTPEVTCVVDVSHSHE 281
QY 69 DPEVKFNHYVDGVEVHNAKTPREEQNSTYRVSVLTVLHQDWLNGKEYKCKVSNKALP 128
Db 282 DPEVKFNHYVDGVEVHNAKTPREEQNSTYRVSVLTVLHQDWLNGKEYKCKVSNKALP 341
QY 129 APIETISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPEN 188
Db 342 APIETISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPEN 401
QY 189 NYKTPPPVLDSDGSFFLYSKLTVDKSRWQOGNVSFSCSVMEALHNNHYTKQSLSPGK 246
Db 402 NYKTPPPVLDSDGSFFLYSKLTVDKSRWQOGNVSFSCSVMEALHNNHYTKQSLSPGK 459

RESULT 15
US-08-957-063-16
; Sequence 16, Application US/08957063
; Patent No. 6025157
;; GENERAL INFORMATION:
;; APPLICANT: Robert D. Klein, Arnon Rosenthal, Mary A. Hynes
;; TITLE OF INVENTION: Neurturin Receptor
;; NUMBER OF SEQUENCES: 19
;; CORRESPONDENCE ADDRESS:

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:43:21 ; Search time 11.3353 Seconds
(without alignments)
1640.982 Million cell updates/sec

Title: 7LINK2

Perfect score: 1334

Sequence: 1 DWLKFYDKVAKLKEAFMD.....MHEALHHVTKLSLSLSPGK 246

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 288829 seqs, 7561385 residues

Total number of hits satisfying chosen parameters: 288829

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pap.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pap.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_PUB.pap.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pap.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pap.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pap.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pap.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pap.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1258	94.3	476	9	US-10-124-905-12
2	1258	94.3	476	9	US-09-948-429B-12
3	1257	94.2	951	9	US-09-935-868-7
4	1254	94.0	859	9	US-09-935-868-7
5	1254	94.0	1158	9	US-09-935-868-26
6	1254	94.0	1168	9	US-09-935-868-24
7	1253	93.9	567	10	US-09-825-561A-16
8	1251	93.8	475	10	US-09-740-002-27
9	1251	93.8	476	10	US-09-747-669-3
10	1250	93.7	399	9	US-09-832-659-2
11	1250	93.7	423	9	US-09-832-659-44
12	1249	93.6	347	9	US-10-091-236-17
13	1249	93.6	347	9	US-10-091-313-7
14	1249	93.6	347	9	US-10-091-268-7
15	1249	93.6	347	10	US-09-796-033-8
16	1249	93.6	347	10	US-09-730-465-8
17	1249	93.6	452	9	US-09-726-258-71
18	1249	93.6	475	10	US-09-740-002-25
19	1249	93.6	547	10	US-09-746-359A-54

20	1249	93.6	571	10	US-09-746-359A-53
21	1249	93.6	594	10	US-09-815-108-22
22	1249	93.6	731	10	US-09-825-012-46
23	1249	93.6	741	10	US-09-825-012-55
24	1248	93.6	418	9	US-09-832-659-42
25	1248	93.6	476	9	US-10-124-905-4
26	1248	93.6	476	9	US-09-948-429B-4
27	1248	93.6	776	9	US-09-935-868-36
28	1248	93.6	776	9	US-09-935-868-40
29	1248	93.6	776	9	US-09-935-868-44
30	1248	93.6	910	9	US-09-935-868-28
31	1246	93.4	380	10	US-09-948-018-39
32	1246	93.4	451	10	US-09-822-698A-26
33	1245	93.3	388	10	US-09-784-623-16
34	1245	93.3	451	9	US-09-925-179-65
35	1245	93.3	451	9	US-09-925-179-66
36	1245	93.3	451	9	US-09-925-179-68
37	1245	93.3	451	10	US-09-920-171-14
38	1245	93.3	451	10	US-09-920-171-16
39	1245	93.3	451	10	US-09-920-171-18
40	1245	93.3	453	9	US-09-925-179-8
41	1245	93.3	453	10	US-09-802-077-8
42	1245	93.3	453	10	US-09-802-096-8
43	1245	93.3	691	9	US-09-935-868-20
44	1245	93.3	694	9	US-09-935-868-22
45	1245	93.3	780	9	US-09-935-868-34

ALIGNMENTS

RESULT 1

US-10-124-905-12
; Sequence 12, Application US/10124905
; Patent No. US20020166136A1

; GENERAL INFORMATION:

; APPLICANT: Anderson, Darrell R.

; TITLE OF INVENTION: "MONKEY MONOCLONAL ANTIBODIES SPECIFIC

; TO HUMAN B7.1 AND/OR B7.2 PRIMATIZED FORMS THEREOF,

; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING, AND USE THEREOF AS

; IMMUNOSUPPRESSANTS"

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS

; STREET: 699 Prince Street

; CITY: Alexandria

; STATE: VA

; COUNTRY: USA

; ZIP: 22314

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10124,905

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/383,916

; FILING DATE:

; APPLICATION NUMBER: US 08/487,550

; FILING DATE: 07-JUN-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Teskin, Robin L.

; REGISTRATION NUMBER: 35,030

; REFERENCE/DOCKET NUMBER: 012712-131

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703-836-6620

; TELEFAX: 703-836-2021

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 476 amino acids

; SEQUENCE CHARACTERISTICS:

; LENGTH: 476 amino acids

```
;
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-10-124-905-12

Query Match          94.3%; Score 1258; DB 9; Length 476;
Best Local Similarity 56.7%; Pred. No. 7.1e-33;
Matches 233; Conservative 5; Mismatches 8; Indels 165; Gaps 1;

QY 1 DWLKAFYD----- 8
   ||: ||
Db 66 EWIGSFSSSGNTYNNPSLKSQVITSTDTSKNOFSLKLSMTAADTAVYYCVDRRLFVSV 125
QY 9 ----- 8
Db 126 GMVYNNWFDVWGPGLVTVSSASTKGPVFPPLAPSSKSTSGGTAALGCLVKDYFPEPTV 185
QY 9 -----KVAEKLK 15
   ||: ||
Db 186 SWNSGALTSQVHTFPVAVLQSSGLYSLSSVTVPSSSILGTQTYICNVNHNKPSNTKVDKAE 245
QY 16 EAFMDKTHTCPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFN 75
Db 246 PKSCDKTHTCPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFN 305
QY 76 WYVDGVEVHNNAKTPREEQYNSTYRVVSVLTVLHQDLNMGKEYCKVSNKALPAPIEKT 135
Db 306 WYVDGVEVHNNAKTPREEQYNSTYRVVSVLTVLHQDLNMGKEYCKVSNKALPAPIEKT 365
QY 136 SKAKGPREPQVYTLPPSRDELTKNOVSLTCLVKGYPSPDIAVEWESNGQPNNTKTPP 195
Db 366 SKAKGPREPQVYTLPPSRDELTKNOVSLTCLVKGYPSPDIAVEWESNGQPNNTKTPP 425
QY 196 VLDSGSEFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKSLSPGK 246
Db 426 VLDSGSEFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKSLSPGK 476

RESULT 2
US-09-948-429B-12
; Sequence 12, Application US/09948429B
; Patent No. US20020177689A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, Darrell R.
; TITLE OF INVENTION: "MONKEY MONOCLONAL ANTIBODIES SPECIFIC
; TO HUMAN B7.1 AND/OR B7.2 PRIMATIZED FORMS THEREOF."
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING, AND USE THEREOF AS
; IMMUNOSUPPRESSANTS"
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: 699 Prince Street
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/948,429B
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/383,916
; FILING DATE:
; APPLICATION NUMBER: US 08/487,550
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
```

```
;
; REFERENCE/DOCKET NUMBER: 012712-131
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6620
; TELEFAX: 703-836-2021
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 476 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-948-429B-12

Query Match          94.3%; Score 1258; DB 9; Length 476;
Best Local Similarity 56.7%; Pred. No. 7.1e-33;
Matches 233; Conservative 5; Mismatches 8; Indels 165; Gaps 1;

QY 1 DWLKAFYD----- 8
   ||: ||
Db 66 EWIGSFSSSGNTYNNPSLKSQVITSTDTSKNOFSLKLSMTAADTAVYYCVDRRLFVSV 125
QY 9 ----- 8
Db 126 GMVYNNWFDVWGPGLVTVSSASTKGPVFPPLAPSSKSTSGGTAALGCLVKDYFPEPTV 185
QY 9 -----KVAEKLK 15
   ||: ||
Db 186 SWNSGALTSQVHTFPVAVLQSSGLYSLSSVTVPSSSILGTQTYICNVNHNKPSNTKVDKAE 245
QY 16 EAFMDKTHTCPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFN 75
Db 246 PKSCDKTHTCPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFN 305
QY 76 WYVDGVEVHNNAKTPREEQYNSTYRVVSVLTVLHQDLNMGKEYCKVSNKALPAPIEKT 135
Db 306 WYVDGVEVHNNAKTPREEQYNSTYRVVSVLTVLHQDLNMGKEYCKVSNKALPAPIEKT 365
QY 136 SKAKGPREPQVYTLPPSRDELTKNOVSLTCLVKGYPSPDIAVEWESNGQPNNTKTPP 195
Db 366 SKAKGPREPQVYTLPPSRDELTKNOVSLTCLVKGYPSPDIAVEWESNGQPNNTKTPP 425
QY 196 VLDSGSEFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKSLSPGK 246
Db 426 VLDSGSEFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKSLSPGK 476

RESULT 3
US-09-935-868-9
; Sequence 9, Application US/09935868
; Patent No. US20020164690A1
; GENERAL INFORMATION:
; APPLICANT: Regeneron Pharmaceuticals, Inc
; TITLE OF INVENTION: Receptor Based Antagonists, and Methods of Making and Using
; FILE REFERENCE: REG 203D
; CURRENT APPLICATION NUMBER: US/09/935,868
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: PCT/US99/22045
; PRIOR FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 951
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-935-868-9

Query Match          94.2%; Score 1257; DB 9; Length 951;
Best Local Similarity 37.0%; Pred. No. 3.7e-32;
Matches 237; Conservative 4; Mismatches 5; Indels 394; Gaps 3;

QY 1 DW----- 2
   ||
Db 312 DMSEASGITYEDRPKAPSFWKIDPSHTQGYRTVOLWKTLPPPEANGKILDYEVTLT 371
```

Qy 3 -----LK 4
Db 372 RWKSHLQNTVNAKLTNLNDRYLATLTVRNLVGKSDAAVLTIPACDFQATHPMDLK 431
Qy 5 AF-----6
Db 432 APPKDNMLWVWTPPRESVKKYILEWCVLSDKAPCIIDMQOEDGTVHRTYLRGNLAESKC 491
Qy 7 -----YD-----8
Db 492 YLITVPVYADPGSPESIKAYLKQAPSGTGTVRTKKVGNKNEAVLEWDQLPVDVQNGFI 551
Qy 9 -----8
Db 552 RNYTIFYRTIIGNETAVNVDSSTHTYTLSSLTSDTLYMVRMAAYTDEGGKDGPEFTFTTP 611
Qy 9 -----8
Db 612 KPAQGEIESGASTKGPSVFPLAPLSKSTSGGTAALGCLVKDYFPEPTVTSNNSGALTSGV 671
Qy 9 -----KVAEKLKEAFMDKTHTCP 26
Db 672 HTFPAVLOSSGLYSLSVVTPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCP 731
Qy 27 PCPAPELLGGPSVFLFPPPKDRLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNA 86
Db 732 PCPAPELLGGPSVFLFPPPKDRLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNA 791
Qy 87 KTKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPAEKTIISKAKGQPREPQ 146
Db 792 KTKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPAEKTIISKAKGQPREPQ 851
Qy 147 VYTLPPSRDELTKNQVSLTCLVKGYGSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLY 206
Db 852 VYTLPPSRDELTKNQVSLTCLVKGYGSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLY 911
Qy 207 SKLTVDKSRWQGNVFSQVMHEALHNHYTQKSLSLSPGK 246
Db 912 SKLTVDKSRWQGNVFSQVMHEALHNHYTQKSLSLSPGK 951

RESULT 4
US-09-935-868-7
; Sequence 7, Application US/09935868
; Patent No. US20020164690A1
; GENERAL INFORMATION:
; APPLICANT: Regeneron Pharmaceuticals, Inc
; TITLE OF INVENTION: Receptor Based Antagonists, and Methods of Making and Using
; FILE REFERENCE: REG 203D
; CURRENT APPLICATION NUMBER: US/09/935.868
; PRIOR FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: PCT/US99/22045
; PRIOR FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 859
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-935-868-7
Query Match 94.0%; Score 1254; DB 9; Length 859;
Best Local Similarity 60.5%; Pred. No. 3.6e-32;
Matches 233; Conservative 8; Mismatches 5; Indels 139; Gaps 3;
Qy 1 DW-----LKAF-----6
Db 469 DWQOEDGTVHRTYLRGNLAESKYLITVTPVYADPGSPESIKAYLKQAPSGTGTVRTK 528
Qy 7 -----YDKVAEKLKEAFM-----19
Db 529 KVGKNEAVLEWDQLPVDVQNGFIRNYTIFYRTIIGNETAVNVDSSTHTYTLSSLTSDTLY 588

Qy 20 -----DKHTCTCPCPAPPELLGGPSVFL 41
Db 589 MVRMAAYTDEGGKDGPEFTFTTPKFAQGEIESGEPKSCDKTHCTCPCPAPPELLGGPSVFL 648
Qy 42 FPPPKDRLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREQYNSTYRV 101
Db 649 FPPPKDRLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREQYNSTYRV 708
Qy 102 VSVLTVLHQDWLNGKEYCKVSNKALPAPAEKTIISKAKGQPREPQVYTLPPSRDELTKNQ 161
Db 709 VSVLTVLHQDWLNGKEYCKVSNKALPAPAEKTIISKAKGQPREPQVYTLPPSRDELTKNQ 768
Qy 162 VSLCLVKGYGSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNV 221
Db 769 VSLCLVKGYGSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNV 828
Qy 222 FSCSVMHEALHNHYTQKSLSLSPGK 246
Db 829 FSCSVMHEALHNHYTQKSLSLSPGK 853
RESULT 5
US-09-935-868-26
; Sequence 26, Application US/09935868
; Patent No. US20020164690A1
; GENERAL INFORMATION:
; APPLICANT: Regeneron Pharmaceuticals, Inc
; TITLE OF INVENTION: Receptor Based Antagonists, and Methods of Making and Using
; FILE REFERENCE: REG 203D
; CURRENT APPLICATION NUMBER: US/09/935.868
; PRIOR FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: PCT/US99/22045
; PRIOR FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
; LENGTH: 1158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-935-868-26
Query Match 94.0%; Score 1254; DB 9; Length 1158;
Best Local Similarity 61.2%; Pred. No. 7.2e-32;
Matches 233; Conservative 8; Mismatches 5; Indels 135; Gaps 3;
Qy 1 DW-----LKAF-----6
Db 778 DWQOEDGTVHRTYLRGNLAESKYLITVTPVYADPGSPESIKAYLKQAPSGTGTVRTK 837
Qy 7 -----YDKVAEKLKEAFM-----19
Db 838 KVGKNEAVLEWDQLPVDVQNGFIRNYTIFYRTIIGNETAVNVDSSTHTYTLSSLTSDTLY 897
Qy 20 -----DKHTCTCPCPAPPELLGGPSVFLFPPK 45
Db 898 MVRMAAYTDEGGKDGPEFTFTTPKFAQGEIESGDKTHCTCPCPAPPELLGGPSVFLFPPK 957
Qy 46 PKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREQYNSTYRVSVL 105
Db 958 PKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREQYNSTYRVSVL 1017
Qy 106 TVLHQDWLNGKEYCKVSNKALPAPAEKTIISKAKGQPREPQVYTLPPSRDELTKNQVSLT 165
Db 1018 TVLHQDWLNGKEYCKVSNKALPAPAEKTIISKAKGQPREPQVYTLPPSRDELTKNQVSLT 1077
Qy 166 CLVKGYGSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFS 225
Db 1078 CLVKGYGSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFS 1137
Qy 226 VMHEALHNHYTQKSLSLSPGK 246
Db 1138 VMHEALHNHYTQKSLSLSPGK 1158

```

RESULT 6
US-09-935-868-24
; Sequence 24, Application US/09935868
; Patent No. US20020164690A1
; GENERAL INFORMATION:
; APPLICANT: Regeneron Pharmaceuticals, Inc
; TITLE OF INVENTION: Receptor Based Antagonists, and Methods of Making and Using
; FILE REFERENCE: REG 203D
; CURRENT APPLICATION NUMBER: US/09/935,868
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: PCT/US99/22045
; PRIOR FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24
; LENGTH: 1168
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-935-868-24

Query Match 94.0%; Score 1254; DB 9; Length 1168;
Best Local Similarity 61.2%; Pred. No. 7.3e-32;
Matches 233; Conservative 8; Mismatches 5; Indels 135; Gaps 3;

QY 1 DW-----LKAF----- 6
||
Db 788 DMOQEDGTVHRTYLRGNLAESKCYLITVTPYADGPGSPESIKAYLKQAPPSKGPTVRTK 847
:|||||:
QY 7 -----YDKVAEKLKEAF----- 19
:|||||:
Db 848 KVGKNEAVLEWDQLPVDVQNGFIRNVTYFIIGNETAVNVDSHTEYTLSSLTSDTLY 907
:|||||:
QY 20 -----DKTHTCPPCPAPPELLGGPSVFLFPPK 45
:|||||:
Db 908 MVRMAAYTDGGKDGPEFTTTPKFAQGETESGDKTHTCPPCPAPPELLGGPSVFLFPPK 967
:|||||:
QY 46 PKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSTYRVVSVL 105
:|||||:
Db 968 PKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSTYRVVSVL 1027
:|||||:
QY 106 TVLHQDLNKGKEYCKVSKNKPAPTEKTSKAKGQPREPQVYTLPPSRDELTKNOVSLT 165
:|||||:
Db 1028 TVLHQDLNKGKEYCKVSKNKPAPTEKTSKAKGQPREPQVYTLPPSRDELTKNOVSLT 1087
:|||||:
QY 166 CLVKGFPSPDIAVWESNGOPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFS 225
:|||||:
Db 1088 CLVKGFPSPDIAVWESNGOPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFS 1147
:|||||:
QY 226 VMHEALHNHYTQKSLSPCK 246
:|||||:
Db 1148 VMHEALHNHYTQKSLSPCK 1168
:|||||:

RESULT 7
US-09-825-561A-16
; Sequence 16, Application US/09825561A
; Patent No. US20020137677A1
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy A.
; APPLICANT: No. US20020137677A1ak, Julia E.
; APPLICANT: West, James W.
; APPLICANT: Presnell, Scott R.
; APPLICANT: Holly, Richard D.
; APPLICANT: Nelson, Andrew J.
; TITLE OF INVENTION: SOLUBLE ZALPHA11 CYTOKINE RECEPTORS
; FILE REFERENCE: 00-22
; CURRENT APPLICATION NUMBER: US/09/825,561A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/194,731
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/222,121
; PRIOR FILING DATE: 2000-07-28

; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: soluble zalphallR/IgGgamma1 polypeptide
US-09-825-561A-16

Query Match 93.9%; Score 1253; DB 10; Length 567;
Best Local Similarity 55.9%; Pred. No. 1.5e-32;
Matches 236; Conservative 3; Mismatches 3; Indels 180; Gaps 3;

QY 5 AFY----- 7
|||
Db 146 AFYMLGKLOYELQYRNRGDPWAVSPRKLISVDSRSVLLPLEFRKDSSEYELQVRAGPM 205
:|||||:
QY 8 -----DKV-----AEKLKEAF----- 18
:|||||:
Db 206 PGSSYQGTWSEWSDPVIFOTQSEELKEGWNPHASTKGPSVFPLAPSSKSTSGGTAALGCL 265
:|||||:
QY 19 ----- 18
:|||||:
Db 266 VKDYFPPEPTVSNWNGALTSVGHVTFPAVLQSSGLYSLSSVVTVPSSSLGTQTYICNVNHHK 325
:|||||:
QY 19 -----MDKTHTCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVD 64
:|||||:
Db 326 PSNTKVDKVKVEPKSCDKTHTCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVD 365
:|||||:
QY 65 VSHEDPEVKFNWYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDLNKGKEYCKVSN 124
:|||||:
Db 386 VSHEDPEVKFNWYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDLNKGKEYCKVSN 445
:|||||:
QY 125 KALPAPIETKTSKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVWESNG 184
:|||||:
Db 446 KALPAPIETKTSKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVWESNG 505
:|||||:
QY 185 QPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFSVMHEALHNHYTQKSLSP 244
:|||||:
Db 506 QPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFSVMHEALHNHYTQKSLSP 565
:|||||:
QY 245 GK 246
:|||||:
Db 566 GK 567
:|||||:

RESULT 8
US-09-740-002-27
; Sequence 27, Application US/09740002
; Patent No. US20020001798A1
; GENERAL INFORMATION:
; APPLICANT: BRAMS, PETER
; APPLICANT: MORROW, PHILLIP
; TITLE OF INVENTION: NEUTRALIZING HIGH AFFINITY HUMAN MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: SPECIFIC TO RSV F-PROTEIN AND METHODS FOR THEIR
; TITLE OF INVENTION: MANUFACTURE AND THERAPEUTIC USE THEREOF
; FILE REFERENCE: 037003-0275759
; CURRENT APPLICATION NUMBER: US/09/740,002
; CURRENT FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 09/335,697
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 08/488,376
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-740-002-27

Query Match 93.8%; Score 1251; DB 10; Length 475;
```

Best Local Similarity 57.2%; Pred. No. 1.2e-32;
Matches 234; Conservative 4; Mismatches 8; Indels 163; Gaps 2;

QY 1 DWL-----KAFYD--- 8
Db 67 EWLARIDWDDTTFYSASLKTRLSISKDTSKNQVVLRTNVPDVTATYFCARASLYDSDS 126
QY 9 ----- 8
Db 127 FYLFYHAYGOGTVTVSSASTKGPSVEPLAPSSKSTSGTAALGCLVKDYFPEPTVSW 186
QY 9 -----KVAEKLKEA 17
Db 187 NSGALTSVHTFPAVLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHPKPSNTKVDKKAEPK 246
QY 18 FMDKTHCPAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVFENWY 77
Db 247 SCDKTHCPAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVFENWY 306
QY 78 VDGVEVHNKTKPREQYNSTYRVSVTLVHODWLNKGKEYCKVSNKALPAPIEKTISK 137
Db 307 VDGVEVHNKTKPREQYNSTYRVSVTLVHODWLNKGKEYCKVSNKALPAPIEKTISK 366
QY 138 AKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVL 197
Db 367 AKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVL 426
QY 198 DSPGSFELYSLVDSRWQGNVFCSCVMHEALHNNHYTKQSLSPGK 246
Db 427 DSPGSFELYSLVDSRWQGNVFCSCVMHEALHNNHYTKQSLSPGK 475

RESULT 9
US-09-747-669-3
; Sequence 3, Application US/09747669
; Patent No. US20020122807A1
; GENERAL INFORMATION:
; APPLICANT: Dan, Michael D.
; APPLICANT: Saleh, Mansoor
; TITLE OF INVENTION: ANTIGEN BINDING FRAGMENTS, DESIGNATED
; TITLE OF INVENTION: 4B5 THAT SPECIFICALLY DETECT CANCER CELLS, NUCLEOTIDES
; TITLE OF INVENTION: ENCODING THE FRAGMENTS, AND USE THEREOF FOR THE PROPHYLAXIS
; TITLE OF INVENTION: AND DETECTION OF CANCERS
; FILE REFERENCE: 316082001001
; CURRENT APPLICATION NUMBER: US/09/747,669
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 09/111,286
; PRIOR FILING DATE: 1998-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic construct
US-09-747-669-3

Query Match 93.8%; Score 1251; DB 10; Length 476;
Best Local Similarity 55.1%; Pred. No. 1.2e-32;
Matches 233; Conservative 6; Mismatches 7; Indels 177; Gaps 2;

QY 1 DWLK-----AFY----- 4
Db 54 NWYRQAPGQGLEWMGNPNKGTGYAKQFGQVTRNTSIRTAYNELSGLRSDTAVY 113
QY 5 -----AFY----- 7
Db 114 FCARNADNVEMAIIHYHGMVNGQGTVTTVSSASTKGPSVFLAPSSKSTSGGTAALGC 173
QY 8 ----- 7
Db 174 LVKDYPPEPVTVSWNSGALTSVGHVTFPAVLQSSGLYSLSSVTVTPSSSLGTQTYICNVNH 233

QY 8 ----DKVAEKLKEAFMDKTHCPAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVW 63
Db 234 KPSNTKVDKVEPKSCDKTHCPAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVW 293
QY 64 DVSHEDPEVFENWYVDGVEVHNKTKPREQYNSTYRVSVTLVHODWLNKGKEYCKV 123
Db 294 DVSHEDPEVFENWYVDGVEVHNKTKPREQYNSTYRVSVTLVHODWLNKGKEYCKV 353
QY 124 NKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESN 183
Db 354 NKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESN 413
QY 184 GQPENNYKTTTPPVLDSDGSFELYSLVDSRWQGNVFCSCVMHEALHNNHYTKQSL 243
Db 414 GQPENNYKTTTPPVLDSDGSFELYSLVDSRWQGNVFCSCVMHEALHNNHYTKQSL 473
QY 244 PGK 246
Db 474 PGK 476

RESULT 10
US-09-832-659-2
; Sequence 2, Application US/09832659
; Patent No. US20020155547A1
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; TITLE OF INVENTION: Interferon-Beta Fusion Proteins and Uses
; FILE REFERENCE: A064PCTSEQ
; CURRENT APPLICATION NUMBER: US/09/832,659
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/120,237
; PRIOR FILING DATE: 1999-02-16
; PRIOR APPLICATION NUMBER: 60/104,491
; PRIOR FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 399
; TYPE: PRT
; ORGANISM: murine
US-09-832-659-2

Query Match 93.7%; Score 1250; DB 9; Length 399;
Best Local Similarity 83.5%; Pred. No. 8.6e-33;
Matches 232; Conservative 5; Mismatches 7; Indels 34; Gaps 1;

QY 3 LKAFYDKVAEKLK-----EAFMDKTHTCPPC 28
Db 122 LKRYGRIHLHLKAKKEYSHCAWTVRVEILRNFYINRLTGYLRNDDDDKVDKTHTCPPC 181
QY 29 PAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNK 88
Db 182 PAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNK 241
QY 89 KPREEQYNSTYRVSVTLVHODWLNKGKEYCKVSNKALPAPIEKTISKAKGQPREQVY 148
Db 242 KPREEQYNSTYRVSVTLVHODWLNKGKEYCKVSNKALPAPIEKTISKAKGQPREQVY 301
QY 149 TLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLDSDGSFFLYSK 208
Db 302 TLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLDSDGSFFLYSK 361
QY 209 LTVDKSRWQGNVFCSCVMHEALHNNHYTKQSLSPGK 246
Db 362 LTVDKSRWQGNVFCSCVMHEALHNNHYTKQSLSPGK 399

RESULT 11
US-09-832-659-44
; Sequence 44, Application US/09832659
; Patent No. US20020155547A1

; APPLICANT: DINGIVAN, CHRISTINE A
; TITLE OF INVENTION: METHODS OF ADMINISTERING/DOSING CD2 ANTAGONISTS FOR THE PREVENTION
; FILE REFERENCE: 10271-054-999
; CURRENT APPLICATION NUMBER: US/10/091,268
; CURRENT FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: US 60/273,098
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/346,918
; PRIOR FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 7
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-091-268-7

Query Match 93.6%; Score 1249; DB 9; Length 347;
Best Local Similarity 80.2%; Pred. No. 6.7e-33;
Matches 235; Conservative 1; Mismatches 8; Indels 49; Gaps 2;

QY 2 WLKAFYDKVAEKLKEAF----- 18
DB 56 WKQK-KDKVAELENSEFRASFKNRVYLDTVSGSLTIYNLTSSDEDEYEMESPNTIDTM 114
QY 19 -----MDKTHTCPPCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 73
DB 115 KFFLYVDKTHTCPPCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 174
QY 74 FNNYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 133
DB 175 FNNYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 234
QY 134 TISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 193
DB 235 TISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 294
QY 194 PPVLSDSGSFFLYSKLTVDKSRWQGNVFSCSVMHEALHNHYTQKSLSLSPGK 246
DB 295 PPVLSDSGSFFLYSKLTVDKSRWQGNVFSCSVMHEALHNHYTQKSLSLSPGK 347

RESULT 15
US-09-796-033-8
; Sequence 8, Application US/09796033
; Patent No. US20020009446A1
; GENERAL INFORMATION:
; APPLICANT: Magilavy, Daniel
; TITLE OF INVENTION: METHOD OF MODULATING MEMORY EFFECTOR
; FILE REFERENCE: 10274-044001
; CURRENT APPLICATION NUMBER: US/09/796,033
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: PCT/US99/20026
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: US 60/098,456
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(28)
US-09-796-033-8

Query Match 93.6%; Score 1249; DB 10; Length 347;
Best Local Similarity 80.2%; Pred. No. 6.7e-33;
Matches 235; Conservative 1; Mismatches 8; Indels 49; Gaps 2;

QY 2 WLKAFYDKVAEKLKEAF----- 18
DB 56 WKQK-KDKVAELENSEFRASFKNRVYLDTVSGSLTIYNLTSSDEDEYEMESPNTIDTM 114
QY 19 -----MDKTHTCPPCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 73
DB 115 KFFLYVDKTHTCPPCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 174
QY 74 FNNYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 133
DB 175 FNNYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 234
QY 134 TISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 193
DB 235 TISKAGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 294
QY 194 PPVLSDSGSFFLYSKLTVDKSRWQGNVFSCSVMHEALHNHYTQKSLSLSPGK 246
DB 295 PPVLSDSGSFFLYSKLTVDKSRWQGNVFSCSVMHEALHNHYTQKSLSLSPGK 347

Search completed: April 21, 2003, 10:51:17
Job time : 13.3353 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:26 ; Search time 18.0882 Seconds
(without alignments)
1307.428 Million cell updates/sec

Title: 7LINK2

Perfect score: 1334

Sequence: 1 DWLKAFYDKVAEKLKEAFMD.....MHEALHNHYTKSLSPGK 246

Scoring table: BLOSUM62

Gapop 10.0 ; Gapext 0.0

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1243	93.2	330	1 GHU	Ig gamma-1 chain C
2	1234	92.5	374	2 S69339	Ig heavy chain V r
3	1233	92.4	255	4 S31866	Ig gamma-1 chain C
4	1190	89.2	234	2 PT0207	Ig gamma chain C r
5	1165	87.3	377	2 A23511	Ig gamma-3 chain C
6	1163	87.2	377	2 A60764	Ig gamma-3 chain C
7	1148	86.1	326	1 G2HU	Ig gamma-2 chain C
8	1144	85.8	327	1 G4HU	Ig gamma-4 chain C
9	1143	85.7	289	1 G3HUM	Ig gamma-4 chain C
10	923	69.2	323	1 GHRB	Ig gamma-3 heavy c
11	917	68.7	328	2 I47160	Ig gamma 2a chain
12	917	68.7	328	2 I47159	Ig gamma 2a chain
13	911	68.3	277	2 I47162	Ig gamma 4 chain c
14	899	67.4	329	1 G2GP	Ig gamma-2 chain C
15	889	66.6	328	2 I47158	Ig gamma-1 chain c
16	885	66.3	328	2 I47161	Ig gamma 3 chain c
17	868	65.1	470	2 S22080	Ig heavy chain pre
18	866	64.9	472	2 S31459	Ig gamma-1 chain -
19	854	64.0	308	2 C30554	Ig heavy chain C r
20	853	63.9	329	1 G3MSC	Ig gamma-3 chain C
21	849	63.6	333	2 PS0018	Ig gamma-2b chain
22	847	63.5	444	2 PC4436	monoclonal antibod
23	842	63.1	398	1 G3MSM	Ig gamma-3 chain C
24	826	61.9	326	2 PS0017	Ig gamma-1 chain C
25	825	61.8	469	2 S37483	Ig gamma-2a chain
26	823	61.7	324	1 G1MS	Ig gamma-1 chain C
27	818	61.3	329	2 S00847	Ig gamma-2c chain
28	818	61.3	393	1 G1MSM	Ig gamma-1 chain C
29	814	61.0	330	1 G2MSA	Ig gamma-2a chain

ALIGNMENTS

RESULT 1

GHU

Ig gamma-1 chain C region - human

C:Species: Homo sapiens (man)

C:Date: 31-Jan-1981 #sequence:Revision 18-Aug-1982 #text_Change 16-Jul-1999

C:Accession: A93433; S36861; S33887; B90563; A90564; B91668; A91723; A02146

R:Ellison, J.W.; Berson, B.J.; Hood, L.E.

Nucleic Acids Res. 10, 4071-4079, 1982

A:Title: The nucleotide sequence of a human immunoglobulin C-gamma1 gene.

A:Reference number: A93433; MUID:82274238; PMID:6287432

A:Accession: A93433

A:Molecule type: DNA

A:Residues: 1-330 <ELL>

A:Cross-references: EMBL:Z17370

A:Note: This sequence has the G1m(17) allotypic marker, 97-Lys, and the G1m(1) marker

A:Note: Lys-330 is removed after translation

R:Harris, L.J.

submitted to the EMBL Data Library, October 1992

A:Reference number: S33904

A:Accession: S36861

A:Molecule type: DNA

A:Residues: 2-330 <HAR>

A:Cross-references: EMBL:Z17370

R:Takahashi, N.; Ueda, S.; Obata, M.; Nikaido, T.; Nakai, S.; Honjo, T.

Cell 29, 671-679, 1982

A:Title: Structure of human immunoglobulin gamma genes: implications for evolution o

A:Reference number: S33887; MUID:83001943; PMID:6811139

A:Accession: S33887

A:Molecule type: DNA

A:Residues: 88-113;235-330 <TAK>

A:Cross-references: EMBL:Z17370

R:Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelm

Biochemistry 9, 3161-3170, 1970

A:Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid se

A:Reference number: A90563; MUID:71064024; PMID:5489771

A:Contents: myeloma protein Eu

A:Accession: B90563

A:Molecule type: protein

A:Residues: 1-96,'R',98-135 <GUN>

A:Note: This sequence has the G1m(3) marker, 97-Arg

R:Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.

Biochemistry 9, 3171-3181, 1970

A:Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid s

A:Reference number: A90564; MUID:71064025; PMID:5530842

A:Contents: Eu

A:Accession: A90564

A:Molecule type: protein

A:Residues: 136-154,'Q',156-165,'Q',167-176,'Q',178-194,'N',196-197,'D',199-238,'E',

A:Note: this sequence has the G1m(non-1) markers, 239-Glu and 241-Met

R:Ponstingl, H.; Hilschmann, N.

Hoppe-Sevler's Z. Physiol. Chem. 357, 1571-1604, 1976

A:Title: Die Primerstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein N

Igen Primaerstruktur.
A:Reference number: A91668; MUID:77070269; PMID:826475
A:Contents: myeloma protein Nie
A:Accession: B91668
A:Molecule type: protein
A:Residues: 1-34, 'Q', '36-96', 'K', '98-115', 'Q', '117-197', 'D', '199-238', 'D', '240', 'L', '242-268', 'E', '273-330' <KHA>
A:Note: this sequence has the Gln(17) and Gln(1) markers
R:Schmidt, W.E.; Jung, H.D.; Palm, W.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 364, 713-747, 1983
A:Title: Die Primaerstruktur des kristallisierten monoklonalen Immunglobulins IgG1 KOL
A:Reference number: A91723; MUID:83289131; PMID:6884994
A:Contents: myeloma protein KOL; disulfide bonds
A:Accession: A91723
A:Molecule type: protein
A:Residues: 1-96, 'R', '98-197', 'D', '199-238', 'E', '240', 'M', '242-266', 'D', '268-271', 'D', '273-330' <SCH>
A:Note: this sequence has the Gln(3) and Gln(non-1) markers
R:Gall, W.E.; Edelman, G.M.
Biochemistry 9, 3188-3196, 1970
A:Title: The covalent structure of a human gammaG-immunoglobulin. X. Intrachain disulfide bonds
A:Reference number: A90565; MUID:71064027; PMID:4923144
A:Contents: annotation; disulfide bonds
R:Dreker, L.; Schwarz, J.; Reichel, W.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 357, 1515-1540, 1976
A:Title: Rule of antibody structure. The primary structure of monoclonal IgG1 immunoglobulin consists of two identical light (kappa) and two identical heavy (lambda) chains. In some cases, such as IgA and IgM, the subunits associate into larger oligomers.
A:Reference number: A91667; MUID:77070267; PMID:1002129
A:Contents: annotation; disulfide bonds
C:Genetics:
A:Gene: GDB:IGHG1
A:Cross-references: GDB:120085; OMIM:147100
A:Map position: 14q32.33-14q32.33
A:Introns: 99/1; 114/1; 224/1
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kappa) and two identical heavy (lambda) chains. In some cases, such as IgA and IgM, the subunits associate into larger oligomers.
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:20-85/Domain: immunoglobulin homology <IM1>
F:137-206/Domain: immunoglobulin homology <IM2>
F:243-310/Domain: immunoglobulin homology <IM3>
F:27-83,144-204,250-308/Disulfide bonds: #status experimental
F:103/Disulfide bonds: interchain (to light chain) #status experimental
F:109,112/Disulfide bonds: interchain (to heavy chain) #status experimental
F:180/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 93.2%; Score 1243; DB 1; Length 330;
Best Local Similarity 96.6%; Pred. No. 5.3e-35;
Matches 230; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 9 KVAEKLKEAFMDKTHRCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHE 68
|||:::|||||
DB 93 KVDKVEPKSCDKTHRCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHE 152
QY 69 DPEVKFNWYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALP 128
|||
DB 153 DPEVKFNWYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALP 212
QY 129 APIETISAKAGPREPOVYTLPPSRDELTKNOVSLTCLVKGFPDSIAVWESNGQCPEN 188
|||
DB 213 APIETISAKAGPREPOVYTLPPSRDELTKNOVSLTCLVKGFPDSIAVWESNGQCPEN 272
QY 189 NYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVYHHEALHNHYTQKSLSLSPGK 246
|||
DB 273 NYKTTTPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVYHHEALHNHYTQKSLSLSPGK 330

RESULT 2
S69339
Ig heavy chain V region precursor - human
C:Species: Homo sapiens (man)
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Dec-2000
C:Accession: S69339; S72664
R:Khamilichi, A.A.; Aucouturier, P.; Preud'homme, J.L.; Cogne, M.
Eur. J. Biochem. 229, 54-60, 1995

A:Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease
A:Reference number: S69339; MUID:95262687; PMID:7744049
A:Accession: S69339
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-374 <KHA>
A:Cross-references: EMBL:X81695
R:Khamilichi, A.A.
submitted to the EMBL Data Library, September 1994
A:Reference number: S72664
A:Accession: S72664
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-140, 'C', '142-374' <KH2>
A:Cross-references: EMBL:X81695
C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 92.5%; Score 1234; DB 2; Length 374;
Best Local Similarity 74.5%; Pred. No. 1.6e-34;
Matches 231; Conservative 6; Mismatches 7; Indels 66; Gaps 3;

QY 1 DWLKAFY---DK----- 9
|||:::|||||
DB 67 EWLALIFWDDDDKRYSPSLRRLTITKDTSKNQVVLTMNVDPADTATYYCGYSVEGYGQG 126
QY 10 -----VAEKLEAFMDKTHRCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTTP 56
|||:::|||||
DB 127 YRFHSMGQGLTVTVSSEPKSC--DKTHRCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTTP 184
QY 57 EVTCVVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGK 116
|||
DB 185 EVTCVVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGK 244
QY 117 EYCKYSNKALPAPIETISAKAGPREPOVYTLPPSRDELTKNOVSLTCLVKGFPDSI 176
|||
DB 245 EYCKYSNKALPAPIETISAKAGPREPOVYTLPPSRDELTKNOVSLTCLVKGFPDSI 304
QY 177 AVEWESNGQPNYKTPPPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVYHHEALHNHYT 236
|||
DB 305 AVEWESNGQPNYKTPPPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVYHHEALHNHYT 364
QY 237 QKSLSLSPGK 246
|||
DB 365 QKSLSLSPGK 374

RESULT 3
S31866
Ig gamma-1 chain C region - synthetic
C:Species: synthetic
C:Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli
C>Date: 06-Jan-1995 #sequence_revision 17-Mar-1997 #text_change 19-May-2000
C:Accession: S31866
R:Filipula, D.
submitted to the EMBL Data Library, February 1993
A:Description: Screening method for protein-protein interactions of cloned gene products
A:Reference number: S31866
A:Accession: S31866
A:Molecule type: mRNA
A:Residues: 1-255 <FIL>
A:Cross-references: EMBL:X70421; NID:g33068; PIDN:CAA49866.1; PID:g33069
C:Keywords: immunoglobulin
F:1-22/Region: Escherichia coli outer membrane protein A precursor
F:23-255/Region: human Ig gamma-1 chain C region

Query Match 92.4%; Score 1233; DB 4; Length 255;
Best Local Similarity 100.0%; Pred. No. 5e-35;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 DKTHRCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 79
|||
DB 29 DKTHRCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 88

Qy 80 GVEVHNAKTKPREQYNSTYRVVSVLTVLHODWLNKKEYCKVSNKALPAPIETISKAK 139
|||||
Db 89 GVEVHNAKTKPREQYNSTYRVVSVLTVLHODWLNKKEYCKVSNKALPAPIETISKAK 148
|||||
Qy 140 GQREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 199
|||||
Db 149 GQREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 208
|||||
Qy 200 DGSFFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNYHTQKSLSPGK 246
|||||
Db 209 DGSFFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNYHTQKSLSPGK 255
|||||

RESULT 4

PT0207
Ig gamma chain C region - chimpanzee
C:Species: Pan troglodytes (chimpanzee)
C:Date: 23-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 16-Jul-1999
R:Ehrlich, P.H.; Moustafa, Z.A.; Oestberg, L.
C:Accession: PT0207
Mol. Immunol. 28, 319-322, 1991
A:Title: Nucleotide sequence of chimpanzee Fc and hinge regions.
A:Reference number: PT0207; MUID:91287716; PMID:2062315
A:Accession: PT0207
A:Molecule type: mRNA
A:Residues: 1-234 <EHR>
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:48-117/Domain: immunoglobulin homology <IMM>

Query Match 89.28; Score 1190; DB 2; Length 234;
Best Local Similarity 95.28; Pred. No. 1.1e-33;
Matches 220; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 9 KVAEKLKEAFMDKTHTCPPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVDVDSHE 68
|| :|||
Db 4 KVDKKVEPKSCDTHTTCCPPCAPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVDVDSHE 63
|||||
Qy 69 DPEVKFNWYVDGVEVHNAKTKPREQYNSTYRVVSVLTVLHODWLNKKEYCKVSNKALP 128
|||||
Db 64 DPEVKFNWYVDGVEVHNAKTKPREQYNSTYRVVSVLTVLHODWLNKKEYCKVSNKALP 123
|||||
Qy 129 APIETISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPEN 188
|||||
Db 124 APIETISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPEN 183
|||||
Qy 189 NYKTTPPVLDSDGSFFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNYHTQKS 239
|||||
Db 184 NYKTTPPVLDSDGSFFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNYHTQKS 234
|||||

RESULT 5

A23511
Ig gamma-3 chain C region (allotype G3m(b)) - human
C:Species: Homo sapiens (man)
C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 23-Jul-1999
R:Huck, S.; Fort, P.; Crawford, D.H.; Lefranc, M.P.; Lefranc, G.
C:Accession: A23511
Nucleic Acids Res. 14, 1779-1789, 1986
A:Title: Sequence of a human immunoglobulin gamma 3 heavy chain constant region gene: cc
A:Reference number: A23511; MUID:86148507; PMID:3081877
A:Accession: A23511
A:Molecule type: DNA
A:Residues: 1-377 <HUC>
A:Cross-references: GB:X03604; GB:M12958; NID:g33070; PIDN:CAA27268.1; PID:g577056
C:Genetics:
A:Gene: IGHG3
A:Cross-references: GDB:119339; OMIM:147120
A:Map position: 14q32.33-14q32.33
A:Introns: 98/3; 115/3; 130/3; 145/3; 160/3; 270/3
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:20-85/Domain: immunoglobulin homology <IMM>

Query Match 87.28; Score 1163; DB 2; Length 377;
Best Local Similarity 76.88; Pred. No. 4.7e-32;
Matches 216; Conservative 8; Mismatches 12; Indels 46; Gaps 2;

RESULT 7

G2HU
Ig gamma-2 chain C region - human
C:Species: Homo sapiens (man)

Query Match 87.3%; Score 1165; DB 2; Length 377;
Best Local Similarity 76.8%; Pred. No. 4e-32;
Matches 218; Conservative 8; Mismatches 12; Indels 46; Gaps 2;

Qy 8 DKVAEKLKEAFMDKTHTCPPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVDVDSHEDEVEQFKWYVDGVE 24
|| :|||
Db 95 DKRYE-LKTPGLGDTHTTCPPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVDVDSHEDEVEQFKWYVDGVE 153
|||||
Qy 25 -CPPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVDVDSHEDEVEQFKWYVDGVE 82
|||||
Db 154 PPCPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVDVDSHEDEVEQFKWYVDGVE 213
|||||
Qy 83 VHNAKTKPREQYNSTYRVVSVLTVLHODWLNKKEYCKVSNKALPAPIETISKAKGQP 142
|||||
Db 214 VHNAKTKPREQYNSTYRVVSVLTVLHODWLNKKEYCKVSNKALPAPIETISKAKGQP 273
|||||
Qy 143 REPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGS 202
|||||
Db 274 REPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGS 333
|||||
Qy 203 FFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNYHTQKSLSPGK 246
|||||
Db 334 FFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNYHTQKSLSPGK 377
|||||

RESULT 6

A60764
Ig gamma-3 chain C region, form LAT - human
C:Species: Homo sapiens (man)
C:Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 16-Jul-1999
C:Accession: A60764
R:Huck, S.; Lefranc, G.; Lefranc, M.P.
Immunogenetics 30, 250-257, 1989
A:Title: A human immunoglobulin IGHG3 allele (Gmb0, b1, c3, c5, u) with an IGHG4 con
A:Reference number: A60764; MUID:90007613; PMID:2571587
A:Accession: A60764
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-377 <HUC>
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:20-85/Domain: immunoglobulin homology <IMM>

Query Match 87.28; Score 1163; DB 2; Length 377;
Best Local Similarity 76.88; Pred. No. 4.7e-32;
Matches 216; Conservative 8; Mismatches 12; Indels 46; Gaps 2;

Qy 8 DKVAEKLKEAFMDKTHTCPPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVDVDSHEDEVEQFKWYVDGVE 24
|| :|||
Db 95 DKRYE-LKTPGLGDTHTTCPPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVDVDSHEDEVEQFKWYVDGVE 153
|||||
Qy 25 -CPPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVDVDSHEDEVEQFKWYVDGVE 82
|||||
Db 154 PPCPCPAPPELLGGPSVFLFPPPKDPTLMISRTPEVTCVVDVDSHEDEVEQFKWYVDGVE 213
|||||
Qy 83 VHNAKTKPREQYNSTYRVVSVLTVLHODWLNKKEYCKVSNKALPAPIETISKAKGQP 142
|||||
Db 214 VHNAKTKPREQYNSTYRVVSVLTVLHODWLNKKEYCKVSNKALPAPIETISKAKGQP 273
|||||
Qy 143 REPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGS 202
|||||
Db 274 REPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGS 333
|||||
Qy 203 FFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNYHTQKSLSPGK 246
|||||
Db 334 FFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNYHTQKSLSPGK 377
|||||


```

||||||| ||: |||||||||||||||||||||||||||||||||||||||
Db 199 YKCVSNKGLPSSTKTSIRAKGQRPQVYTLPPSQEEMTKNOVSLTCLVKGFPSDIA 258
QY 178 VEWESNGOPENNYKTTTPVLDSGDSFYLKSLTVDKSRWQGNVFSVMHEALHNHYTQ 237
||||||| ||: |||||||||||||||||||||||||||||||||||||||
Db 259 VEWESNGOPENNYKTTTPVLDSGDSGFFLYSLTVDKSRWQGNVFSVMHEALHNHYTQ 318
QY 238 KSLSLSPGK 246
||||||| ||: |||||||||||||||||||||||||||||||||||||||
Db 319 KSLSLSLGK 327

RESULT 9
G3HUW1
Ig gamma-3 heavy chain disease proteins - human
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1979 #sequence_revision 23-Oct-1981 #text_change 16-Jul-1999
C:Accession: A90442; A92219; A90198; A93915; A02149
R:Frangione, B.; Rosenwasser, E.; Prelli, F.; Franklin, E.C.
Biochemistry 19, 4304-4308, 1980
A:Title: Primary structure of human gamma3 immunoglobulin deletion mutant: gamma3 heavy-
A:Reference number: A90442; MUID:81021548; PMID:6774747
A:Contents: heavy chain disease protein wis
A:Accession: A90442
A:Molecule type: protein
A:Residues: 1-289 <FRA>
A:Note: the molecule is a dimer linked by 12 disulfide bonds; it has an extra interchain
A:Note: this protein lacks most of the V region and all of the CH1 region. Residue 12 co
A:Note: the sequence of residues 42-76 was taken from the reference that follows
R:Michaelson, T.E.; Frangione, B.; Franklin, E.C.
J. Biol. Chem. 252, 883-889, 1977
A:Title: Primary structure of the 'hinge' region of human IgG3. Probable quadruplication
A:Reference number: A92219; MUID:77118561; PMID:402363
A:Contents: normal gamma-3 chains, sequence corresponding to residues 12-97 of protein w
A:Accession: A92219
A:Molecule type: protein
A:Residues: 12-97 <MIC>
A:Note: the hinge region in gamma-3 chains is about four times as long as in other gamma
idue segment (12-28)
R:Wolfenstein-Rodei, C.; Frangione, B.; Prelli, F.; Franklin, E.C.
Biochem. Biophys. Res. Commun. 71, 907-914, 1976
A:Title: The amino acid sequence of "heavy chain disease" protein ZUC. Structure of the
A:Reference number: A90198; MUID:77021516; PMID:823945
A:Contents: heavy chain disease protein Zuc, partial sequence corresponding to residues
A:Accession: A90198
A:Molecule type: protein
A:Residues: 59-125, 'EB', 128-226, 228-289 <WOL>
A:Note: this protein lacks most of the V region, all of the CH1 region, and part of the
R:Alexander, A.; Steinmetz, M.; Barritault, D.; Frangione, B.; Franklin, E.C.; Hood, L.;
Proc. Natl. Acad. Sci. U.S.A. 79, 3260-3264, 1982
A:Title: gamma heavy chain disease in man: cDNA sequence supports partial gene deletion
A:Reference number: A93915; MUID:82247835; PMID:6808505
A:Contents: heavy chain disease protein Omn
A:Accession: A93915
A:Molecule type: mRNA
A:Residues: 12-70; 72-114; 116-125, 'E', 127-133, 'L', 135-136, 'E', 138, 'Y', 140-154, 'D', 156-157
A:Note: a carboxyl-terminal Lys is removed posttranslationally
A:Note: this sequence may represent an allelic form or another gamma chain subclass
C:Comment: The heavy chain disease protein wis is shown.
C:Genetics:
A:Gene: GDB:IGHG3
A:Cross-references: GDB:119339; OMIM:147120
A:Map position: 14q32.33-14q32.33
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; immunoglobulin; pyroglutamic acid
F:203-270/Domain: immunoglobulin homology <IMM>
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:6.140/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 85.7%; Score 1143; DB 1; Length 289;
Best Local Similarity 74.7%; Pred. No. 9.5e-32;
Matches 210; Conservative 15; Mismatches 11; Indels 45; Gaps 1;

```

```

QY 10 VAEKLKEAFMDKTHT----- 24
||: |||
Db 9 VSSELKTLPLGDTHTCPRCPEPKSCDTPPCPCPEKSCDTPPCPCPEKSCDTPPP 68
QY 25 CPCPCAPPELLGGPSVFLFPKPKDITLMSRTPEVTVVVDVSHEDPEYKFNWYDGVVH 84
||: |||
Db 69 CPCPCAPPELLGGPSVFLFPKPKDITLMSRTPEVTVVVDVSHEDPEYKFNWYDGVVH 128
QY 85 NAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPRE 144
||||||| ||: |||||||||||||||||||||||||||||||||||||||
Db 129 NAKTKPREQQFNSTFRVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPRE 188
QY 145 PQVYTLPPSRDELTKNOVSLTCLVKGFVPSDIAVEWESNGQPENNYKTPPVLDSDGSFF 204
||||||| ||: |||||||||||||||||||||||||||||||||||||||
Db 189 PQVYTLPPSRDEMTKNOVSLTCLVKGFVPSDIAVEWESNGQPENNYKTPPVLDSDGSFF 248
QY 205 LYSKLTVDKSRWQGNVFSVMHEALHNHYTQKSLSLSPG 245
||||||| ||: |||||||||||||||||||||||||||||||||||||||
Db 249 LYSKLTVDKSRWQGNVFSVMHEALHNHYTQKSLSLSPG 289

RESULT 10
GHRB
Ig gamma chain C region - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 24-Apr-1984 #sequence_revision 15-Nov-1984 #text_change 16-Jul-1999
C:Accession: A91749; A90290; A93928; A90245; A94416; A02161
R:Bernstein, K.E.; Alexander, C.B.; Mage, R.G.
Immunogenetics 18, 387-397, 1983
A:Title: Nucleotide sequence of a rabbit IgG heavy chain from the recombinant F-I ha;
A:Reference number: A91749; MUID:84030930; PMID:6313520
A:Accession: A91749
A:Molecule type: mRNA
A:Residues: 1-323 <BER>
A:Note: this sequence has the d12 allotypic marker, 104-Thr, and the e14 marker, 185
R:Pratt, D.M.; Mole, L.E.
Biochem. J. 151, 337-349, 1975
A:Title: Sequence studies on the constant region of the Fd sections of rabbit immuno
A:Reference number: A90290; MUID:76135469; PMID:1243651
A:Accession: A90290
A:Molecule type: protein
A:Residues: 1-47, 'E', 49-71, 'PV', 72-128 <PRA>
R:Martens, C.L.; Moore, K.W.; Steinmetz, M.; Hood, L.; Knight, K.L.
Proc. Natl. Acad. Sci. U.S.A. 79, 6018-6022, 1982
A:Title: Heavy chain genes of rabbit IgG; isolation of a cDNA encoding gamma heavy c
A:Reference number: A93928; MUID:83299917; PMID:6193512
A:Accession: A93928
A:Molecule type: mRNA
A:Residues: 88-103, 'M', 105-143, 'E', 145-184, 'A', 186, 'E', 188-266 <MAR>
A:Cross-references: GB:M16426; NID:g165111; PIDN:AAA31289.1; PID:g165112
A:Note: this sequence has the d11 allotypic marker, 104-Met, and the e15 allotypic m
R:Fruchter, R.G.; Jackson, S.A.; Mole, L.E.; Porter, R.R.
Biochem. J. 116, 249-259, 1970
A:Title: Sequence studies of the Fd section of the heavy chain of rabbit immunoglobul
A:Reference number: A90245; MUID:70110015; PMID:5461106
A:Accession: A90245
A:Molecule type: protein
A:Residues: 132-143, 'E', 145-161 <FRU>
R:Hill, R.B.; Lebovitz, H.E.; Fellows Jr., R.E.; Delaney, R.
in Gamma Globulins, Nobel Symp. 3, Killander, J., ed., pp.109-127, Almqvist and Wiks
A:Reference number: A94416
A:Accession: A94416
A:Molecule type: protein
A:Residues: 129-131; 155-172, 'D', 174-184, 'A', 186, 'E', 188-200, 'D', 202-217, 'E', 219-232, '
A:Note: this has the e15 allotypic marker, 185-Ala
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate int
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:20-82/Domain: immunoglobulin homology <IM1>
F:130-199/Domain: immunoglobulin homology <IM2>
F:236-303/Domain: immunoglobulin homology <IM3>

```



```

||||||| 1:|||||:||||||| 1:|||||:||||||| 1:|||||:||||||| 1:|||||:|||||||
Db 88 TCVVVDVSDQENPEVQSWYVDGVEVHTAQTARKPEEQENSTYRVVSVLPVLPQHODWLNCKEF 147

QY 119 KCKVSNKALPAPIEKTITSKAKGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAV 178
||||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:|||||
Db 148 KCKVNNKDLPAPIITRIILSKAKGQTPQVYTLPPPTTELSRSKVTLCVTGTFYPPDIV 207

QY 179 EWESNGO--PENNYKTTTPPVLDSDGSEFLYSLKLTVDKSRWQGNVFCVSNVHEALHNYHT 236
||||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:|||||
Db 208 EMQWNGQPEPGNRYRTTPPOQVDGTYFLYSLKLAVDKASQWGRGDTFOCAVWHEALHNYHT 267

QY 237 QKSLSLSPGK 246
||||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:|||||
Db 268 QKSIFKTPGK 277

```

RESULT 14

G2GP

```

Ig gamma-2 chain C region - guinea pig
C:Species: Cavia porcellus (guinea pig)
C:Date: 07-May-1981 #sequence_revision 07-May-1981 #text_change 16-Jul-1999
C:Accession: A94553; A90352; A90359; A90384; A90385; A02151
R:Trischmann, T.M.
Submitted to the Atlas, April 1975
A:Reference number: A94553
A:Accession: A94553
A:Molecule type: protein
A:Residues: 1-3 <TRI>
R:Birshtein, B.K.; Hussain, O.Z.; Cebra, J.J.
Biochemistry 10, 18-25, 1971
A:Title: Structure of heavy chain from strain 13 guinea pig immunoglobulin-G(2). III. An
A:Reference number: A90352; MUID:71058471; PMID:5538606
A:Accession: A90352
A:Molecule type: protein
A:Residues: 4-68 <BIR>
R:Turner, K.J.; Cebra, J.J.
Biochemistry 10, 9-17, 1971
A:Title: Structure of heavy chain from strain 13 guinea pig immunoglobulin-G(2). II. An
A:Reference number: A90359; MUID:71058486; PMID:5538616
A:Accession: A90359
A:Molecule type: protein
A:Residues: 69-133;312-329 <TUR>
R:Tracey, D.E.; Cebra, J.J.
Biochemistry 13, 4796-4803, 1974
A:Title: Primary structure of the C-H2 homology region from guinea pig IgG2 antibodies.
A:Reference number: A90384; MUID:75036072; PMID:4429665
A:Accession: A90384
A:Molecule type: protein
A:Residues: 134-226 <TRA>
R:Trischmann, T.M.; Cebra, J.J.
Biochemistry 13, 4804-4811, 1974
A:Title: Primary structure of the C-H3 homology region from guinea pig IgG2 antibodies.
A:Reference number: A90385; MUID:75036073; PMID:4609467
A:Accession: A90385
A:Molecule type: protein
A:Residues: 227-311 <TR2>
R:Oliveira, B.; Lamm, M.E.
Biochemistry 10, 26-31, 1971
A:Title: Interchain disulfide bridges of guinea pig gamma-2- immunoglobulin.
A:Reference number: A90354; MUID:71058474; PMID:4922544
A:Contents: annotation; disulfide bonds
A:Note: Cys-16 is involved in a heavy-light chain bond
A:Note: Cys-105, Cys-107, and Cys-110 form inter-heavy chain bonds
C:Comment: This chain was isolated from pooled serum of strain 13 inbred guinea pigs.
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (k)
chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into 14
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:21-81/Domain: immunoglobulin homology <IM1>
F:135-204/Domain: immunoglobulin homology <IM2>
F:241-310/Domain: immunoglobulin homology <IM3>
F:28-79/Disulfide bonds: #status experimental
F:142-202/Disulfide bonds: #status experimental

```

```

F:178/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:248-308/Disulfide bonds: #status experimental

Query Match 67.4%; Score 899; DB 1; Length 329;
Best Local Similarity 67.1%; Pred. No. 3.7e-23;
Matches 167; Conservative 25; Mismatches 44; Indels 13; Gaps 2;

QY 10 VAERLKEAFMDKT-----HTCPGPCAPPELLGGPSVFLFPPKPKDTLMISRTPEV 58
||||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:|||||
Db 81 VAHPASSTKVDKTVETPTP2PBECTCPKCPPEENLGGPSVFIEPPKPKDTLMISLTPRV 140

QY 59 TCVVVDVSHEDPEVKFNMYVDGVEVHNAKTPREEQYNSTYRVVSVLTVHLQDWLNGKEY 118
||||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:|||||
Db 141 TCVVVDVSDQENPEVQSWYVDGVEVHTAQTARKPEEQENSTYRVVSVLPVLPQHODWLNCKEF 200

QY 119 KCKVSNKALPAPIEKTITSKAKGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAV 178
||||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:|||||
Db 201 KCKVNNKDLPAPIEKTITSKAKGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAV 260

QY 179 EWESNGO--PENNYKTTTPPVLDSDGSEFLYSLKLTVDKSRWQGNVFCVSNVHEALHNYHT 236
||||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:|||||
Db 261 EWASNRVPVSEKEKYKNTPTTEDADGSEFLYSLKLTVDKSRWQGNVFCVSNVHEALHNYHT 320

QY 237 QKSLSLSPGK 245
||||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:|||||
Db 321 QKAISRSPGK 329

RESULT 15
I47158
Ig gamma 1 chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47158
R:Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of
A:Reference number: I47158; MUID:95015845; PMID:7930579
A:Accession: I47158
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-328 <KAC>
A:Cross-references: EMBL:U03778; NID:G433121; PIDN:AAA52216.1; PID:G433122
C:Gene: IgG1
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 66.6%; Score 889; DB 2; Length 328;
Best Local Similarity 70.1%; Pred. No. 8.2e-23;
Matches 164; Conservative 29; Mismatches 38; Indels 3; Gaps 2;

QY 15 KEAFMDKTHCTCPGPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKF 74
||||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:|||||
Db 96 KRVGIHQPTCTPICPGCE-VAGPSVFIFPPKPKDTLMISQTPETVTCVVVDVSKGHAETQF 154

QY 75 NWYVDGVEVHNAKTPREEQYNSTYRVVSVLTVHLQDWLNGKEYCKVSNKALPAPIEKT 134
||||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:|||||
Db 155 SWYVDGVEVHTAETPRKEQFNSTYRVVSVLPVLPQHODWLNCKEFPKPKVNNVLDLPAPITRT 214

QY 135 ISKAKGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGO--PENNYKT 192
||||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:|||||
Db 215 ISKAIGOSRPPQVYTLPPPAEELSRSKVTLCVLIGYFPDPIHVEWKSNGQPEPENTYRT 274

QY 193 TTPVLDSDGSEFLYSLKLTVDKSRWQGNVFCVSNVHEALHNYHTQKLSLSPGK 246
||||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:||||| 1:|||||
Db 275 TTPQODVDGTFYFLYSLKLAVDKARWDGDKFECAVWHEALHNYHTQKLSLSPGK 328

```

Search completed: April 21, 2003, 10:49:18

Job time : 19.0882 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:21 ; Search time 9.40588 seconds
(without alignments)
1084.766 Million cell updates/sec

Title: 7LINK2

Perfect score: 1334

Sequence: 1 DWLKAFYDKVAELKEAFMD.....MHEALHNYTKSLSLSPCK 246

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1243	93.2	330	GC1_HUMAN	P01857 homo sapien
2	1148	86.1	290	GC3_HUMAN	P01860 homo sapien
3	1148	86.1	326	GC2_HUMAN	P01859 homo sapien
4	1144	85.8	327	GC4_HUMAN	P01861 homo sapien
5	923	69.2	323	GC_RABIT	P01870 oryctolagus
6	899	67.4	329	GC2_CAVPO	P01862 cavia porce
7	853	63.9	329	GC3_MOUSE	P22436 mus musculus
8	849	63.6	333	GC3_MOUSE	P20761 rattus norv
9	842	63.1	398	GC3_MOUSE	P03987 mus musculus
10	826	61.9	326	GC1_RAT	P20759 rattus norv
11	823	61.7	324	GC1_MOUSE	P01868 mus musculus
12	818	61.3	329	GC3_MOUSE	P20762 rattus norv
13	818	61.3	393	GC1_MOUSE	P01869 mus musculus
14	814	61.0	330	GC3_MOUSE	P01863 mus musculus
15	809	60.6	335	GC3_MOUSE	P01864 mus musculus
16	809	60.6	399	GC3_MOUSE	P01865 mus musculus
17	795	59.6	322	GC3_MOUSE	P20760 rattus norv
18	789	59.1	336	GC3_MOUSE	P01866 mus musculus
19	784	58.8	405	GC3_MOUSE	P01867 mus musculus
20	388	29.1	421	EPC_MOUSE	P06336 mus musculus
21	381	28.6	454	MUC_MOUSE	P01871 homo sapien
22	378	28.3	429	EPC_MOUSE	P01855 rattus norv
23	377	28.3	391	MUC_MOUSE	P04220 homo sapien
24	377	28.3	455	MUC_MOUSE	P01872 mus musculus
25	376	28.2	428	EPC_MOUSE	P01854 homo sapien
26	373	28.0	458	MUC_MOUSE	P03988 oryctolagus
27	367	27.5	476	MUC_MOUSE	P01873 mus musculus
28	363	27.2	479	MUC_MOUSE	P04221 oryctolagus
29	356	26.7	454	MUC_MOUSE	P06337 mesocricetu
30	351	26.3	450	MUC_MOUSE	P20768 suncus muri
31	349	26.2	450	MUC_MOUSE	P01874 canis fami
32	334	25.0	438	HVC2_HETFR	P23085 heterodontu
33	332	24.9	438	HVC2_HETFR	P23087 heterodontu

ALIGNMENTS

```

RESULT 1
GC1_HUMAN          STANDARD;          PRT;          330 AA.
AC P01857:
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig gamma-1 chain C region.
GN IGHL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82274238; PubMed=6287432;
RA Ellison J.W., Berson B.J., Hood L.E.;
RT "The nucleotide sequence of a human immunoglobulin C gamma gene.";
RL Nucleic Acids Res. 10:4071-4079(1982).
RN [2]
RP SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).
RX MEDLINE=71064024; PubMed=5489771;
RA Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,
RA Waxdal M.J., Edelman G.M.;
RT "The covalent structure of a human gamma G-immunoglobulin. VII. Amino
RL acid sequence of heavy-chain cyanogen bromide fragments H1-H4.";
RN Biochemistry 9:3161-3170(1970).
RP SEQUENCE OF 136-329 (EU).
RX MEDLINE=71064025; PubMed=5530842;
RA Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,
RA Edelman G.M.;
RT "The covalent structure of a human gamma G-immunoglobulin. 8. Amino
RL acid sequence of heavy-chain cyanogen bromide fragments H5-H7.";
RN Biochemistry 9:3171-3181(1970).
RP SEQUENCE (MYELOMA PROTEIN NIE).
RX MEDLINE=77070269; PubMed=826475;
RA Ponstingl H., Hilschmann N.;
RT "The rule of antibody structure. The primary structure of a
RL monoclonal IgG1 immunoglobulin (myeloma protein NIE). III. The
RL chymotryptic peptides of the H-chain, alignment of the tryptic
RL peptides and discussion of the complete structure.";
RN Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).
RP SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.
RX MEDLINE=83289131; PubMed=6884994;
RA Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;
RT "Three-dimensional structure determination of antibodies. Primary
RL structure of crystallized monoclonal immunoglobulin IgG1 KOL, I.";
RN Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).
RP DISULFIDE BONDS.
RX MEDLINE=71064027; PubMed=4923144;
RA Gall W.E., Edelman G.M.;
RT "The covalent structure of a human gamma G-immunoglobulin. X.

```

```

34 326 24.4 446 1 MUC_CHICK
35 324 24.3 461 1 HVC2_HETFR
36 321 24.1 393 1 HVC3_HETFR
37 315 23.6 370 1 HVC1_HETFR
38 312 23.4 6669 1 NEBU_HUMAN
39 311 23.3 6486 1 TYCC_BACBR
40 310 23.2 299 1 ALC_RABIT
41 308 23.1 6359 1 BACC_BACLI
42 306 22.9 4393 1 PGSM_HUMAN
43 303 22.7 4969 1 RYR2_RABIT
44 302 22.6 353 1 RYR1_GORGO
45 302 22.6 4829 1 BIR6_HUMAN

```

```

P01875 gallus gall
P23088 heterodontu
P23086 heterodontu
P23084 heterodontu
P20929 homo sapien
O30409 b tyrocidin
P01879 oryctolagus
O68008 b bacitraci
P98160 homo sapien
P30957 oryctolagus
P20758 gorilla gor
Q9n909 homo sapien

```

Intrachain disulfide bonds."; Biochemistry 9:3188-3196(1970). [7]

DISULFIDE BONDS.

RX MEDLINE=77070267; PubMed=1002129;

RA Dreker L., Schwarz J., Reichel W., Hilschmann N.;

RT "Rule of antibody structure. The primary structure of a monoclonal

RT IgG1 immunoglobulin (myeloma protein Nie), I: Purification and

RT characterization of the protein, the L- and H-chains, the

RT cyanogen bromide cleavage products, and the disulfide bridges.";

RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).

[8]

RX X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).

RP MEDLINE=81208100; PubMed=7236608;

RA Daisenhofer J.;

RT "Crystallographic refinement and atomic models of a human Fc fragment

RT and its complex with fragment B of protein A from Staphylococcus

RT aureus at 2.9- and 2.8-A resolution.";

RL Biochemistry 20:2361-2370(1981).

CC -1- MISCELLANEOUS: NIE HAS THE GIM(17) ALLOTYPIC MARKER, 97-K, & THE

CC GIM(1) MARKERS, 239-D & 241-L. KOL & EU SEQUENCES HAVE THE GIM(3)

CC MARKER & THE GIM (NON-1) MARKERS.

CC -1- MISCELLANEOUS: NIE ALSO DIFFERS IN THE AMIDATION STATES OF

CC 35,116,198,269 & 272.

CC -1- MISCELLANEOUS: EU ALSO DIFFERS IN THE AMIDATION STATES OF RESIDUES

CC 155, 166, 177, 195, 198, 269, AND 272 AND IN THE ORDER OF RESIDUES

CC 268-272.

CC -1- MISCELLANEOUS: KOL ALSO DIFFERS IN THE AMIDATION STATES OF

CC RESIDUES 198,267&272.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; J00228; AAC82527.1; ALT_INIT.

DR PTR; A02146; GHHU.

DR PDB; 1FC1; 15-JUL-92.

DR PDB; 1FC2; 15-JUL-92.

DR Genew; HGNC:5525; IGHG1.

DR MLM; 147100; -.

DR InterPro; IPR003006; Ig_MHC.

DR InterPro; IPR003597; Ig_cl.

DR InterPro; IPR003600; Ig_1like.

DR Pfam; PF00047; Ig; 3.

DR SMART; SM00410; Ig_1like; 1.

DR SMART; SM00407; IGcl; 2.

DR PROSITE; PS00290; Ig_MHC; 2.

KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;

KW 3D-structure.

FT NON_TER 1 1

FT DOMAIN 1 98 CHI.

FT DOMAIN 99 110 HINGE.

FT DOMAIN 111 223 CH2.

FT DOMAIN 224 330 CH3.

FT DISULFID 27 83

FT DISULFID 103 103 INTERCHAIN (WITH LIGHT CHAIN).

FT DISULFID 109 109 INTERCHAIN (WITH HEAVY CHAIN).

FT DISULFID 112 112 INTERCHAIN (WITH HEAVY CHAIN).

FT DISULFID 144 204

FT DISULFID 250 308

FT CARBOHYD 180 180

FT MOD_RES 330 330

FT VARIANT 97 97

FT VARIANT 239 239

FT VARIANT 241 241

FT STRAND 123 126

N-LINKED (GLCNAC. . .).

REMOVED POST-TRANSLATIONALLY.

K -> R (IN GIM(3) MARKER).

/FTId=VAR_003886.

D -> E (IN GIM(NON-1) MARKER).

/FTId=VAR_003887.

L -> M (IN GIM(NON-1) MARKER).

/FTId=VAR_003888.

[2]
 RN REVISIONS TO 12-97 (PROTEIN WIS).
 RX MEDLINE=77118561; PubMed=402363;
 RA Michaelson T.E., Frangione B., Franklin E.C.;
 RT "Primary structure of the 'hinge' region of human IgG3. Probable
 RL quadruplication of a 15-amino acid residue basic unit.";
 J. Biol. Chem. 252:883-889(1977).
 [3]
 RN REVISIONS TO 59-289 (PROTEIN WIS/DISEASE PROTEIN ZUC).
 RX MEDLINE=77021516; PubMed=823945;
 RA Wolfenstein-Todel C., Frangione B., Prelli F., Franklin E.C.;
 RT "The amino acid sequence of 'heavy chain disease' protein ZUC.
 RL Structure of the FC fragment of immunoglobulin G3.";
 Biochem. Biophys. Res. Commun. 71:907-914(1976).
 [4]
 RN SEQUENCE FROM N.A. (DISEASE PROTEIN OMM).
 RX MEDLINE=82247835; PubMed=6808505;
 RA Alexander A., Steinmetz M., Barriault D., Frangione B.,
 RA Franklin E.C., Hood L., Buxbaum J.N.;
 RT "Gamma Heavy chain disease in man: cDNA sequence supports partial
 RL gene deletion model.";
 Proc. Natl. Acad. Sci. U.S.A. 79:3260-3264(1982).
 CC -1- SUBUNIT: DIMER LINKED BY 12 DISULFIDE BONDS; IT HAS AN EXTRA
 CC INTERCHAIN DISULFIDE BOND AT POSITION 7 IN ADDITION TO THE 11
 CC NORMALLY PRESENT IN THE HINGE REGION.
 CC -1- MISCELLANEOUS: THE HEAVY CHAIN DISEASE PROTEIN WIS IS SHOWN.
 CC REF. 2
 CC -1- MISCELLANEOUS: DISEASE PROTEIN WIS IS LACKING MOST OF THE V REGION
 CC AND ALL OF THE CH1 REGION.
 CC -1- MISCELLANEOUS: DISEASE PROTEIN ZUC LACK MOST OF THE V REGION, ALL
 CC OF THE CH1 REGION, AND PART OF THE HINGE COMPARED WITH NORMAL
 CC GAMMA-3 HEAVY CHAINS.
 CC -1- MISCELLANEOUS: DISEASE PROTEIN OMM MAY REPRESENT AN ALLELIC FORM
 CC OR ANOTHER GAMMA CHAIN SUBCLASS.
 CC -1- MISCELLANEOUS: THE HINGE REGION IN GAMMA-3 CHAINS IS ABOUT FOUR
 CC TIMES AS LONG AS IN OTHER GAMMA CHAINS AND CONTAINS THREE
 CC IDENTICAL 15-RESIDUE SEGMENTS PRECEDED BY A SIMILAR 17-RESIDUE
 CC SEGMENT (12-28).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; J00231; AAA52805.1; ALT_SEQ.
 DR PIR; A02149; G3HUW1.
 DR HSP; P01857; 1FC1.
 DR Genew; HGNC:5527; IGHG3.
 DR MIM; 147120; .
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003597; Ig-cl.
 DR InterPro; IPR003600; Ig_like.
 DR Pfam; PF00047; ig; 2.
 DR SMART; SM00410; Ig_like; 1.
 DR SMART; SM00407; IGcl; 1.
 DR PROSITE; PS00290; Ig_MHC; 1.
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein; Repeat.
 FT DOMAIN 12 73
 FT HINGE.
 FT DOMAIN 74 183
 FT CH2.
 FT DOMAIN 184 289
 FT CH3.
 FT REPEAT 29 43
 FT REPEAT 44 58
 FT REPEAT 59 73
 FT MOD_RES 1 1
 FT CARBOHYD 6 6
 FT N-LINKED (GLNAC. . .).
 FT INTERCHAIN (WITH HEAVY CHAIN DIMER).
 FT DISULFID 7 7
 FT INTERCHAIN (WITH HEAVY CHAIN DIMER).
 FT DISULFID 24 24
 FT INTERCHAIN (WITH HEAVY CHAIN DIMER).
 FT DISULFID 27 27
 FT INTERCHAIN (WITH HEAVY CHAIN DIMER).
 FT DISULFID 33 33
 FT INTERCHAIN (WITH HEAVY CHAIN DIMER).

FT DISULFID 39 39 INTERCHAIN (WITH HEAVY CHAIN DIMER).
 FT DISULFID 42 42 INTERCHAIN (WITH HEAVY CHAIN DIMER).
 FT DISULFID 48 48 INTERCHAIN (WITH HEAVY CHAIN DIMER).
 FT DISULFID 54 54 INTERCHAIN (WITH HEAVY CHAIN DIMER).
 FT DISULFID 57 57 INTERCHAIN (WITH HEAVY CHAIN DIMER).
 FT DISULFID 63 63 INTERCHAIN (WITH HEAVY CHAIN DIMER).
 FT DISULFID 69 69 INTERCHAIN (WITH HEAVY CHAIN DIMER).
 FT DISULFID 72 72 INTERCHAIN (WITH HEAVY CHAIN DIMER).
 FT CARBOHYD 140 140 N-LINKED (GLNAC. . .).
 FT MOD_RES 290 290 REMOVED POST-TRANSLATIONALLY.
 FT VARIANT 126 127 QV -> EB (IN ZUC).
 FT VARIANT 134 134 /FTID=VAR_003890.
 FT VARIANT 139 139 P -> L (IN OMM).
 FT VARIANT 182 182 F -> Y (IN OMM).
 FT VARIANT 227 227 /FTID=VAR_003892.
 FT VARIANT 227 227 T -> A (IN OMM).
 FT VARIANT 227 227 S -> N (IN OMM).
 FT VARIANT 227 227 /FTID=VAR_003893.
 FT VARIANT 227 227 /FTID=VAR_003894.
 FT VARIANT 279 279 MISSING (IN ZUC).
 FT VARIANT 279 279 /FTID=VAR_003895.
 FT VARIANT 279 279 F -> Y (IN OMM).
 FT VARIANT 279 279 /FTID=VAR_003896.
 SQ SEQUENCE 290 AA; 32331 MW; E69CB95705B2F46 CRC64;
 Query Match 86.1%; Score 1148; DB 1; Length 290;
 Best Local Similarity 74.8%; Pred. No. 2.le-32;
 Matches 211; Conservative 15; Mismatches 11; Indels 45; Gaps 1;
 QY 10 VAELKELAFMDKTHRT----- 24
 I: : : : :
 Db 9 VSSELKPLGDTTHTCPCPEPKSCDTPPCPCPEPKSCDTPPCPCPEPKSCDTPPP 68
 QY 25 CPCCPAPPELLGGPSVFLFPPPKPDKTLMISRTPEVTVVVDVSHEDPEVKFNNYVGVH 84
 I: : : : :
 Db 69 CPCCPAPPELLGGPSVFLFPPPKPDKTLMISRTPEVTVVVDVSHEDPEVKFNNYVGVH 128
 QY 85 NAKTKPREOYNRYVSVLTVLHODWLNKGKVKYKSNKALPAPIETKTSKAGOPRE 144
 I: : : : :
 Db 129 NAKTKPREOYNRYVSVLTVLHODWLNKGKVKYKSNKALPAPIETKTSKAGOPRE 188
 QY 145 PQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSF 204
 I: : : : :
 Db 189 PQVYTLPPSRDEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSF 248
 QY 205 LYSKLTVDKSRWQQGNVPCFVSCVMEALHNHYTQKSLSLSPGK 246
 I: : : : :
 Db 249 LYSKLTVDKSRWQQGNVPCFVSCVMEALHNHYTQKSLSLSPGK 290
 RESULT 3
 GC2_HUMAN
 ID GC2_HUMAN STANDARD; PRT; 326 AA.
 AC P01859;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ig gamma-2 chain C region.
 GN IGHG2.
 OS Homo sapiens (Human).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBITaxID=9606;
 [1]
 RN SEQUENCE OF 2-326 FROM N.A.
 RX MEDLINE=82197621; PubMed=6804948;
 RA Ellison J.W., Hood L.E.;
 RT "Linkage and sequence homology of two human immunoglobulin gamma
 RL heavy chain constant region genes.";
 Proc. Natl. Acad. Sci. U.S.A. 79:1984-1988(1982).
 RN [2]
 RP SEQUENCE OF 88-115 FROM N.A.


```

RP SEQUENCE OF 1-30 AND 81-326.
RX MEDLINE=70207560; PubMed=4192699;
RA Pink J.R.L., Buttery S.H., de Vries G.M., Millstein C.;
RT "Human immunoglobulin subclasses. Partial amino acid sequence of the
RL constant region of a gamma 4 chain.";
RL Biochem. J. 117:33-47(1970).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K01316; AAB59394.1; ALT_INIT.
DR PIR; A02150; G4HU.
DR HSP; P01842; 7FAB.
DR Genew; HGNC:5528; IGHG4.
DR MIM; 147130; -.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; Ig_Like; 1.
DR SMART; SM00407; IG_C1; 2.
DR PROSITE; PS00290; IG_MHC; 2.
KW Immunoglobulin domain; Immunoglobulin C region.
FT DOMAIN 1 98
FT CH1.
FT DOMAIN 99 110
FT HINGE.
FT DOMAIN 111 220
FT CH2.
FT DOMAIN 221 327
FT CH3.
FT DISULFID 14 14
FT INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 83
FT DISULFID 106 106
FT DISULFID 109 109
FT DISULFID 141 201
FT DISULFID 247 305
SQ SEQUENCE 327 AA; 35940 MW; 3EBDB811EF208E7A CRC64;

Query Match 85.8%; Score 1144; DB 1; Length 327;
Best Local Similarity 84.7%; Pred. No. 4.1e-32;
Matches 211; Conservative 9; Mismatches 6; Indels 23; Gaps 1;

QY 21 KTHT-----CPCPAPELLGGPSVFLFPPKPKDTLMISRTPE 57
Db 79 KTYCNVDHKPSNWKVKRVESKYGPPCPAPEFLGGPSVFLFPPKPKDTLMISRTPE 138
QY 58 VTCVVYDVSHEDPEVAFNNYVDGVEVHNKTPREQYNSTYRVSVLTVLHODWLNKGE 117
Db 139 VTCVVYDVSDQDEPEVQFNYYVDGVEVHNKTPREQFNSTYRVSVLTVLHODWLNKGE 198
QY 118 YKCKVSNKALPAPIETKISKAKQPREPVYTLPPSDELTKNQVSLTCLVKGFYPSDIA 177
Db 199 YKCKVSNKGLPSSIEKISKAKQPREPVYTLPPSDEEMTKNQVSLTCLVKGFYPSDIA 258
QY 178 VEVESNGQPNNTKTPPVLDSDGSPFLYSLKLTVDKSRWQGNVFCSVNHEALHNYHQ 237
Db 259 VEVESNGQPNNTKTPPVLDSDGSPFLYSLKLTVDKSRWQGNVFCSVNHEALHNYHQ 318
QY 238 KSLSLSPGK 246
Db 319 KSLSLSLGK 327

RESULT 5
GC_RABIT STANDARD; PRT; 323 AA.
AC P01870;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)

```

```

DE Ig gamma chain C region.
OS Eryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
ON NCBI_TaxID=9986;
RX SEQUENCE FROM N.A.
RX MEDLINE=84030930; PubMed=6313520;
RA Bernstein K.E., Alexander C.B., Mage R.G.;
RT "Nucleotide sequence of a rabbit IgG heavy chain from the recombinant
RT F-1 haplotype.";
RL Immunogenetics 18:387-397(1983).
RN [2]
RP SEQUENCE OF 1-128.
RX MEDLINE=76135469; PubMed=1243651;
RA Pratt D.M., Mole L.E.;
RT "Sequence studies on the constant region of the Fd sections of rabbit
RT immunoglobulin G of different allotype.";
RL Biochem. J. 151:337-349(1975).
RN [3]
RP SEQUENCE OF 88-266 FROM N.A.
RX MEDLINE=83299917; PubMed=6193512;
RA Martens C.L., Moore K.W., Steinmetz M., Hood L., Knight K.L.;
RT "Heavy chain genes of rabbit IgG: isolation of a cDNA encoding gamma
RT heavy chain and identification of two genomic C gamma genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:6018-6022(1982).
RN [4]
RP SEQUENCE OF 132-161.
RX MEDLINE=70110015; PubMed=5461106;
RA Fruchter R.G., Jackson S.A., Mole L.E., Porter R.R.;
RT "Sequence studies of the Fd section of the heavy chain of rabbit
RT immunoglobulin G.";
RL Biochem. J. 116:249-259(1970).
RN [5]
RP SEQUENCE OF 129-131 AND 155-322.
RA Hill R.L., Lebovitz H.E., Fellows R.E. Jr., Delaney R.;
RL (In) Killander J. (eds.);
RL Gamma globulins, Nobel symp. 3, pp.109-127, Almqvist and Wiksell,
RL Stockholm (1967).
CC -!- MISCELLANEOUS: REF.1 SEQUENCE HAS THE D12 ALLOTYPE MARKER,
CC 104-THR, AND THE E14 MARKER, 185-THR. REF.3 HAS THE D11 AND E15
CC MARKERS AND REF.5 THE E15 MARKER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M16426; AAA31289.1; -.
DR PIR; A02161; GHRB.
DR HSP; P01857; 1FC1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_C1.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00407; IG_C1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT VARIANT 104 104
FT T -> M (IN D11 MARKER).
FT VARIANT 185 185
FT T -> A (IN E15 MARKER).
FT CONFLICT 48 48
FT N -> E (IN REF. 2).
FT CONFLICT 71 71
FT V -> VPV (IN REF. 2).
FT CONFLICT 144 144
FT Q -> E (IN REF. 3 AND 4).
FT CONFLICT 173 173
FT N -> D (IN REF. 5).
FT CONFLICT 187 187
FT Q -> E (IN REF. 3 AND 5).
FT CONFLICT 201 201
FT N -> D (IN REF. 5).
FT CONFLICT 218 218
FT Q -> E (IN REF. 5).
FT CONFLICT 233 233
FT E -> Q (IN REF. 5).
FT CONFLICT 246 246
FT N -> D (IN REF. 5).
FT CONFLICT 256 256
FT E -> G (IN REF. 5).

```

```

FT CONFLICT 260 260 N -> D (IN REF. 5).
FT CONFLICT 266 266 N -> D (IN REF. 5).
FT CONFLICT 280 280 Y -> W (IN REF. 5).
FT CONFLICT 284 284 N -> S (IN REF. 5).
SQ SEQUENCE 323 AA; 35404 MW; 69E8A118D579A8B CRC64;

Query Match 69.2%; Score 923; DB 1; Length 323;
Best Local Similarity 69.8%; Pred. No. 1.7e-24;
Matches 169; Conservative 29; Mismatches 39; Indels 5; Gaps 2;

QY 10 VAEKLKEAFMDKT---HTC--PPCAPPELLGGPSVFLPPPKPDTLMISRTPEVTCVVVD 64
DB 82 VAHPATNTKVTAPSTCKTPPELGGPSVFLPPPKPDTLMISRTPEVTCVVVD 141
QY 65 VSHDEPEVKFNMYVDGVEVHNKTRPREQYNSYTVRVSVLVHLQDNLNGREYKCKVSN 124
DB 142 VSQDDPEVQFTWYINNEQVTRAPPLREQQFNSTIRVVSTLTITQDNLGRKFKCKVHN 201
QY 125 KALPAPIETISKAGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNG 184
DB 202 KALPAPIETISKAGQPLEKPVYTMGPPEELSSRSVSLTCMNGFYPSDISVEWENKNG 261
QY 185 OPENNYKTTTPVLDSGSPFLSKYLVTKVDSKRWQGNVFCSCVMHEALHNNHYTKQSLSLSP 244
DB 262 KAEDNTKTTTPAVLDSDGSYFLYLNKLSVPTSEWQRGDVFCTCSVMHEALHNNHYTKQSLRSP 321
QY 245 GK 246
DB 322 GK 323

RESULT 6
GC2_CAVPO STANDARD; PRT; 329 AA.
AC P01862;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2 chain C region.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE OF 1-3.
RA Trischmann T.M.;
RL Submitted (APR-1975) to the PIR data bank.
RN [2]
RP SEQUENCE OF 4-68.
RX MEDLINE=71058471; PubMed=5538606;
RA Birshstein B.K.; Hussain Q.Z.; Cebra J.J.;
RT "Structure of heavy chain from strain 13 guinea pig
RT immunoglobulin-G(2). 3. Amino acid sequence of the region around the
RT half-cystine joining heavy and light chains.";
RL Biochemistry 10:18-25(1971).
RN [3]
RP SEQUENCE OF 69-133 AND 312-329.
RX MEDLINE=71058486; PubMed=5538616;
RA Turner K.J.; Cebra J.J.;
RT "Structure of heavy chain from strain 13 guinea pig
RT immunoglobulin-G(2). II. Amino acid sequence of the carboxyl-terminal
RT and hinge region cyanogen bromide fragments.";
RL Biochemistry 10:9-17(1971).
RN [4]
RP SEQUENCE OF 134-226.
RX MEDLINE=75036072; PubMed=4429665;
RA Tracey D.E.; Cebra J.J.;
RT "Primary structure of the CH2 homology region from guinea pig IgG2
RT antibodies.";
RL Biochemistry 13:4796-4803(1974).
RN [5]
RP SEQUENCE OF 227-311.
RX MEDLINE=75036073; PubMed=4609467;

```

```

RA Trischmann T.M.; Cebra J.J.;
RT "Primary structure of the CH3 homology region from guinea pig IgG2
RT antibodies.";
RL Biochemistry 13:4804-4811(1974).
RN [6]
RP DISULFIDE BONDS.
RX MEDLINE=71058474; PubMed=4922544;
RA Oliveira B.; Lamm M.E.;
RT "Interchain disulfide bridges of guinea pig gamma-2-immunoglobulin.";
RL Biochemistry 10:26-31(1971).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM POOLED SERUM OF STRAIN
CC 13 INBRED GUINEA PIGS.
DR PIR; A02151; G2GP.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00410; Ig_Like; 1.
DR SMART; SM00407; IGC1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NOX_TER 1 1
FT DISULFID 16 16 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 28 79
FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 142 202
FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .).
FT DISULFID 248 308
SQ SEQUENCE 329 AA; 36074 MW; 5D231B7164D1FBA9 CRC64;

Query Match 67.4%; Score 899; DB 1; Length 329;
Best Local Similarity 67.1%; Pred. No. 1.2e-23;
Matches 167; Conservative 25; Mismatches 44; Indels 13; Gaps 2;

QY 10 VAEKLKEAFMDKT-----HTCPPCAPPELLGGPSVFLPPPKPDTLMISRTPEV 58
DB 81 VAHPASSYTKVDKTVETPTP2PBCTCPKCPENLGGPSVFLPPPKPDTLMISLTPRV 140
QY 59 TCVVVVDVSHEDPEVKFNMYVDGVEVHNKTRPREQYNSYTVRVSVLVHLQDNLNGREY 118
DB 141 TCVVVVDVSDPEVQFTWFDNKPVGNAETKPRVQYNTTFRVESVLPITQHDNLGRKEF 200
QY 119 KCKVSNKALPAPIETISKAGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAV 178
DB 201 KCKVYNKALPAPIETISKAGAPRMPDVYTLPPSRDELTSKSVSVTCLINFFPADIHV 260
QY 179 EWESNGQP--ENNYYKTTTPVLDSGSPFLSKYLVTKVDSKRWQGNVFCSCVMHEALHNNHYT 236
DB 261 EWASNRVPVSEKEYKNTPTIEDAGSYFLYSLTVDKSAWDQGTVYTCVSMHEALHNNHYT 320
QY 237 QKSLSLSPG 245
DB 321 QKAISRSPG 329

RESULT 7
GC3_MOUSE
ID GC3_MOUSE STANDARD; PRT; 329 AA.
AC P22436;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig gamma-3 chain C region, secreted form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;

```



```

RA Wels J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBO J. 3:2041-2046(1984).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00451; ; NOT_ANNOTATED_CDS.
DR PIR; B02156; G3MSC.
DR HSP; P01857; IFCL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig-cl.
DR Pfam; PF000360; Ig_like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; Ig_like; 1.
DR SMART; SM00407; IGcl; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
SQ SEQUENCE 329 AA; 36228 MW; F45827174182BAD6 CRC64;

Query Match 63.9%; Score 853; DB 1; Length 329;
Best Local Similarity 55.8%; Pred. No. 4.7e-22;
Matches 155; Conservative 37; Mismatches 45; Indels 41; Gaps 3;

QY 10 VAELKKEAFM-----DKTHTC-----PP--C 28
Db 52 VSSVLQSGFYSLSLVTPSPSTWPSQTVICNAHPASKTELKRIEPRKPSPPGSSC 111
QY 29 PAPELIGGSPVFLPPPKDTLMISRTPEVTCVVDVSHEDPEVKFWYVDGVEVHNAKT 88
Db 112 PPGNIGLGGSPVFIIPPKPKDLMISLTPKVTVCVVDVSEDDPDVHVSFVDNKEVHTAWT 171
QY 89 KPREEQNSYRVVSVLTVLHQDLNGKVKCKVSNKALPAPIEKTISKAKGQPREPQVY 148
Db 172 QPRAQVNSYRVVSVLPIQHODMRCGEKCKVNNKALPAPIERTISKPKGRAQTQVY 231
QY 149 TLPPSRDELTKNOVSLTCLVKGYFSPVDIAVWESNGOPENNYKTPPVLDSDGFFLYSK 208
Db 232 TIPPPEQMSKKVSLTCLVTNPFSEISVWEWERNGELEQDKNTPTPILSDGTYFLYSK 291
QY 209 LTVDKSRWQGNVFCSCVMHEALHNHTQKLSLSPK 246
Db 292 LTVDTDSWLOGEIFTCSVWHEALHNHTQKLSRSPK 329

RESULT 8
GCB_RAT
ID GCB_RAT STANDARD; PRT; 333 AA.
AC G20761;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2B chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;

"Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
RT PIR; PS0018; PS0018.
DR HSP; P01842; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig-cl.
DR Pfam; PF000360; Ig_like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; Ig_like; 1.
DR SMART; SM00407; IGcl; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 80
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 147 207
FT DISULFID 253 311
SQ SEQUENCE 333 AA; 36497 MW; 55F8B64D48D460A6 CRC64;

Query Match 63.6%; Score 849; DB 1; Length 333;
Best Local Similarity 63.2%; Pred. No. 6.8e-22;
Matches 153; Conservative 34; Mismatches 52; Indels 3; Gaps 1;

QY 8 DKVAELKEAFMDKTHTCP---PCPAPELIGGSPVFLPPPKDTLMISRTPEVTCVVD 64
Db 92 DKVRRNGGIGHKCPTCTCHKCPVPELLGGSPVFIIPPKPKDILLISQNAKVTGVVD 151
QY 65 VSHEDPEVKFNWYVDGVEVHNAKTPREEQNSYRVVSVLTVLHQDLNGKVKCKVSN 124
Db 152 VSEEDPDVQSNFVNNVEVHTAQTPREEQNSYRVVSVLPIQHODMRCGEKCKVNN 211
QY 125 KALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYFSPVDIAVWESNG 184
Db 212 KALPSPIEKTISKPKGLVRKPVVYMGPPTEBOLTEQVSLTCLTSGFLPDIGVWTSNG 271
QY 185 QPENNYKTPPVLDSDGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHTQKLSLSP 244
Db 272 HIEKNYKTEPVMSDGSFFMYSKLVNRSRWDSPAPFCVSVHGLHNHHVKSISRPP 331
QY 245 GK 246
Db 332 GK 333

RESULT 9
GCB_MOUSE
ID GCB_MOUSE STANDARD; PRT; 398 AA.
AC P03987;
DT 23-OCT-1986 (Rel. 02, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-3 chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;
RA Wels J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBO J. 3:2041-2046(1984).
RN [2]
RP SEQUENCE OF 328-398 FROM N.A.
RX MEDLINE=8401483; PubMed=6314258;
RA Komaromy M., Clayton L., Rogers J., Robertson S., Kettman J.,
RA Wall R.;
RT "The structure of the mouse immunoglobulin in gamma 3 membrane gene

```

```

RT segment.";
RL Nucleic Acids Res. 11:6775-6785(1983).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J00451; AAB59655.1; -.
DR EMBL: V01526; CAA24767.1; ALT_SEQ.
DR PIR: A02155; G3MSM.
DR HSP: P01857; 1FC1.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_C1.
DR Pfam: PF00047; Ig_3.
DR SMART: SM00410; Ig_Like; 1.
DR SMART: SM00407; IGcl; 2.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
FT TRANSMEM 346 362 POTENTIAL.
FT DOMAIN 363 398 CYTOPLASMIC (POTENTIAL).
FT CONFLICT 333 338 E -> G (IN REF. 2).
FT CONFLICT 342 342 E -> F (IN REF. 2).
FT CONFLICT 388 388 P -> Q (IN REF. 2).
SQ SEQUENCE 398 AA; 43929 MW; C7F264B50A41B95 CRC64;

Query Match 63.1%; Score 842; DB 1; Length 398;
Best Local Similarity 51.3%; Pred. No. 2e-21;
Matches 154; Conservative 41; Mismatches 47; Indels 58; Gaps 4;

QY 3 LKAFYDK-----VAEKLKEAFM-----DKTHTC----- 25
DB 28 VAGIEPEPVTKWNGALSSGVRTVSSVLQSGFYSLSLVTPVSPSTWPSQTVCINVAHPA 87
QY 26 -----PP-----CPAPELLGGPSVFLPPPKPKDTLMISRTPEVTCVVVD 64
DB 88 SKTELKRIEPRIPKSTPPGSSCPGNILGSPSVFIIPPCKDALMSLPKVTCTVVVD 147
QY 65 VSHEDPEVKFNWYDGVGVHNAKTPREEQNSTYRVVSVLTVLHQQDLNKGKCKVSN 124
DB 148 VSEDDPDVHVSFWFVONKEVHTAQTQPREAQYNSTPRVVSALPIQHQQDMRGKFKCKVNN 207
QY 125 KALPAPIETKISKAGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEMESNG 184
DB 208 KALPAPIETKISKAGRAQTPOVYTIIPPREQMSKKVSLTCLVTFNFFSEALISVEWERNG 267
QY 185 QPENNYKTPPVLDSDGFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNHYTQKSLSLSP 244
DB 268 ELEQDYKNTPTPLDSDGTGYFLYSKLTVDTSNLOGEITFCVSVVHEALHNHHTQKNLSRSP 327

RESULT 10
GCL_RAT
ID GCL_RAT STANDARD; PRT; 326 AA.
AC P20759;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 36, Last annotation update)
DE Ig gamma-1 chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;

```

```

RN SEQUENCE FROM N.A.
RX MEDLINE-89232738; PubMed-3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
DR PIR: PS0017; PS0017.
DR HSP: P01842; 7FAB.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_C1.
DR Pfam: PF00047; Ig_3.
DR SMART: SM00407; IGcl; 2.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 112 HINGE.
FT DOMAIN 113 219 CH2.
FT DOMAIN 220 326 CH3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 140 200
FT DISULFID 246 304
FT CARBOHYD 176 176 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 326 AA; 35946 MW; 013BAB45EF49B9DA CRC64;

Query Match 61.9%; Score 826; DB 1; Length 326;
Best Local Similarity 59.1%; Pred. No. 4e-21;
Matches 146; Conservative 42; Mismatches 45; Indels 14; Gaps 3;

QY 10 VAEKLKEAFMOKT-----HTCPCPAPELLGG---PSVFLPPPKPKDTLMISRTPEVT 59
DB 84 VAHPASSTKVDKIVPRNCGGCKPC----ICTGEVSSVFIPPKPKDVLITITPKVT 139
QY 60 CVVVDVSHEDPEVKFNWYDGVGVHNAKTPREEQNSTYRVVSVLTVLHQQDLNKGKVK 119
DB 140 CVVVDISQDDPEVHFSEWFDVVEHTAQTTPREEQNSTERSVSELPILHQQDLNNGRTFR 199
QY 120 CKVSNKALPAPIETKISKAGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVE 179
DB 200 CKVTSAAAPSPETKISKPEGRTOVPHYVTMSPTKEEMTQNEVITCVKGFYPPDIYE 259
QY 180 WESNGQPNYKTPPVLDSDGFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNHYTOKS 239
DB 260 WQMGQPOENYKNTPTTMDTDSGYFLYSKLVKVKKEWQGNFTTCSVLHEGLHNHHTKS 319
QY 240 LSLSPGK 246
DB 320 LSHSPGK 326

RESULT 11
GCL_MOUSE
ID GCL_MOUSE STANDARD; PRT; 324 AA.
AC P01868;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ig gamma-1 chain C region.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN SEQUENCE FROM N.A.
RX MEDLINE-80045036; PubMed=115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
gamma 1 chain gene.";

```

```

Cell 18:559-568(1979).
[2]
SEQUENCE OF 76-324 FROM N.A. (MYELOMA PROTEIN MOPC 31C).
RX MEDLINE=80202559; PubMed=6769752;
RA Obata M., Yamawaki-Kataoka Y., Takahashi N., Kataoka T., Shimizu A.,
RA Mano Y., Selman J.G., Peterlin B.M., Leder P., Honjo T.;
RT "Immunoglobulin gamma 1 heavy chain gene: structural gene sequences
RT cloned in a bacterial plasmid.";
RL Gene 9:87-97(1980).
RL [3]
RP SEQUENCE OF 70-322 FROM N.A. (MYELOMA PROTEIN MOPC 21).
RX MEDLINE=80012837; PubMed=113776;
RA Rogers J., Clarke P., Salser W.;
RT "Sequence analysis of cloned cDNA encoding part of an immunoglobulin
RT heavy chain.";
RL Nucleic Acids Res. 6:3305-3321(1979).
[4]
RP SEQUENCE (MYELOMA PROTEIN MOPC 21).
RX MEDLINE=78242288; PubMed=98524;
RA Adetugbo K.;
RT "Evolution of immunoglobulin subclasses. Primary structure of a
RT murine myeloma gammal chain.";
RL J. Biol. Chem. 253:6068-6075(1978).
[5]
RP DISULFIDE BONDS (MOPC 21).
RX MEDLINE=73008889; PubMed=5073237;
RA Svasti J., Milstein C.;
RT "The disulphide bridges of a mouse immunoglobulin G1 protein.";
RL Biochem. J. 126:837-850(1972).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00793; CAA24172.1; -
DR EMBL; V00793; CAA24173.1; -
DR EMBL; V00793; CAA24174.1; -
DR EMBL; V00793; CAA24175.1; -
DR EMBL; V00795; CAA24176.1; -
DR PIR; A02159; GIMS.
DR HSSP; P01842; 7FAB.
DR GlycoSuiteDB; P01868; -
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; Igcl; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing.
FT NON_TER 1 1
FT DOMAIN 1 97 CHI.
FT DOMAIN 98 110 HINGE.
FT DOMAIN 111 217 CH2.
FT DOMAIN 218 324 CH3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 138 198
FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .).
/FTID=CAR_000055.
FT FT 244 302
FT DISULFID 244 302
FT MOD_RES 324 324 REMOVED POST-TRANSLATIONALLY.
FT CONFLICT 276 276 N -> D (IN REF. 3).
FT CONFLICT 278 278 N -> D (IN REF. 3).
SQ SEQUENCE 324 AA; 35704 MW; A338812F3D1F2C93 CRC64;

```

```

Query Match 61.7%; Score 823; DB 1; Length 324;
Best Local Similarity 59.7%; Pred. No. 4.9e-21;
Matches 141; Conservative 48; Mismatches 42; Indels 5; Gaps 2;

QY 13 KLKEAFMDKTHTCPP--CPAPELLGGPSVFLFPKPKDITLMISRTPEVTCVVDVSHEDP 70
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 92 KVDKKIVPRDCGCKPGICTVPEV---SSVFIFPKPKDVLITITLPKVTCCVVDISKDDP 148
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 71 EVKFNMYVDGVEVHNKTKPREQYNSTYRVSVLTVLHODWNLGKEYCKVSNKALPAP 130
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 149 EVQFSWFDVDDVEVHTAQTQPREEQNFTRSVSELPFHHQDNLGKFKCRVNSAAPAP 208
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 131 IEKTIKAKQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNY 190
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 209 IEKTIKTKGRAPQVYTIPTPKPEQMAKDKVSLTCHITDFFPEDITVEQWNGQPAENY 268
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 191 KTTPPVLDSDGSFLLSKLTVDKSRWQOGNVFSGSVMEALHNHYTKSLSPGK 246
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 269 KNTQPIMNTNGSYFVYSKLVNQRKNWEAGNTFTCSVLHGLHNHHTKSLSHSPGK 324

RESULT 12
GCC_RAT GCC_RAT STANDARD; PRT; 329 AA.
ID AC P20762;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2C chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88166903; PubMed=3127222;
RA Brueggemann M., Delmastro-Galfre P., Waldmann H., Calabi F.;
RT "Sequence of a rat immunoglobulin gamma 2c heavy chain constant
RT region cDNA: extensive homology to mouse gamma 3.";
RL Eur. J. Immunol. 18:317-319(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X07189; CAA30169.1; -
DR PIR; S00847; S00847.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00410; Ig_like; 1.
DR SMART; SM00407; Igcl; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DOMAIN 1 97 CHI.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 222 CH2.
FT DOMAIN 223 329 CH3.
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 113 113 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 143 203
FT DISULFID 249 307
SQ SEQUENCE 329 AA; 36571 MW; 5FCD7B7933850773 CRC64;

```

```
Query Match 61.3%; Score 818; DB 1; Length 329;
Best Local Similarity 57.9%; Pred. No. 7.7e-21;
Matches 146; Conservative 42; Mismatches 38; Indels 26; Gaps 2;

QY 21 KTHC-----PP-----CPAPELLGPGSVFLPPPKDITLMSR 54
DB 78 QTVCVVAHPATKSNLIKRIEPRPKPRPTDICSDDNGLRPSVFIFPPKPKDITLMTL 137
QY 55 TPEVTCVVVDVSHEDPEVKFNKYVDGVEVHNKATKPREOYNSTYRVVSVLTVLDHQLN 114
DB 138 TRKVCVVVDVSEEDPDVGFQSVFNDVRVFTAQTPQHHEQLNGTKRVVSTLHIQHDWMS 197
QY 115 GREYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPS 174
DB 198 GREYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPS 257
QY 175 DTAWEESNGQPNKYKTPPVLDSGGFFLYSKLTVDKSRWQQGNVSCSVMHEALNHH 234
DB 258 SISVWERNGELEDQYKNTLPVLDSDESYFLYSLKSLVDTSWMRGDIKTSVYVHEALNHH 317
QY 235 YTKSLSLSPGK 246
DB 318 HQKNLSRSPGK 329

RESULT 13
GC1M_MOUSE STANDARD; PRT; 393 AA.
AC P01869.
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ig gamma-1 chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=80045036; PubMed=115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT Cloning and complete nucleotide sequence of mouse immunoglobulin
RT gamma 1 chain gene.;
RL Cell 18:559-568(1979).
RN [2]
RP SEQUENCE OF 323-393 FROM N.A.
RX MEDLINE=82197626; PubMed=6804950;
RA Tyler B.M., Cowman A.F., Gerondakis S.D., Adams J.M., Bernard O.;
RT "mRNA for surface immunoglobulin gamma chains encodes a highly
RT conserved transmembrane sequence and a 28-residue intracellular
RT domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2008-2012(1982).
RN [3]
RP SEQUENCE OF 323-366 FROM N.A.
RX MEDLINE=82115295; PubMed=6799207;
RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,
RA Eisenberg D., Wall R.;
RT "Gene segments encoding transmembrane carboxyl termini of
RT immunoglobulin gamma chains.";
RL Cell 26:19-27(1981).
RN [4]
RP SEQUENCE OF 1-44 FROM N.A.
RX MEDLINE=8222190; PubMed=6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
RT immunoglobulin gamma chains.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
CC -!- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA
CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
CC BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
CC IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND

CC SEGMENT OF MU CHAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; V00793; CAA24172.1; -
CC EMBL; V00793; CAA24173.1; -
CC EMBL; V00793; CAA24174.1; -
CC PIR; B02159; GLMSM.
CC HSP; P01842; 7FAB.
CC MGD; MGI:96446; Igh-4.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR003597; Ig_c1.
CC Pfam; PF00047; Ig; 3.
CC SMART; SM00407; Igel; 2.
CC PROSITE; PS00290; IG_MHC; 1.
CC Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
CC Alternative splicing; Transmembrane.
CC NON_TER 1
CC DOMAIN 1 97 CH1.
CC DOMAIN 98 110 HINGE.
CC DOMAIN 111 217 CH2.
CC DOMAIN 218 324 CH3.
CC DISULFID 27 82
CC DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).
CC DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
CC DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
CC DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
CC DISULFID 138 198 N-LINKED (GLCNAC. . .).
CC CARBOHYD 174 174 POTENTIAL.
CC DISULFID 244 302 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 340 357 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 358 393
CC SEQUENCE 393 AA; 43386 MW; 4CC88343B7A1CE27 CRC64;

Query Match 61.3%; Score 818; DB 1; Length 393;
Best Local Similarity 59.6%; Pred. No. 1.3e-20;
Matches 140; Conservative 48; Mismatches 42; Indels 5; Gaps 2;

QY 13 KLKEAFMDKTKTCTPP--CPAPELLGPGSVFLPPPKDITLMSRTPETVCVVVDVSHEDP 70
DB 92 KVDKKIVPRDCGCKPCICTVPEV---SSVFIPPKPKDVLITITPKVTCVVVDISKDDP 148
QY 71 EVKFNWYVDGVEVHNKATKPREOYNSTYRVVSVLTVLDHQLNGKEYKCKVSNKALPAP 130
DB 149 EVQFSWFVDYDEVHTAQTPREOYNSTFRSVSELPIMHQDLNGKKEKFCRVNSAAPPAP 208
QY 131 IEKTIKRAKGPQREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNY 190
DB 209 IEKTIKTRKGRKAPQVYTIIPPKKEQMAKDKVSLTCTMITDFFEDITVEMQMGQPAENY 268
QY 191 KTTTPPVLDGSGFFLYSKLTVDKSRWQQGNVSCSVMHEALNHHYTKSLSLSPG 245
DB 269 KNTQPIMNNGSYFYVSKLVNQSKEWAGNTFTCSVLHEGLNHHHTEKSLSHSPG 323

RESULT 14
GCAA_MOUSE STANDARD; PRT; 330 AA.
ID GCAA_MOUSE
AC P01863;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ig gamma-2A chain C region, A allele.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
```

1] SEQUENCE FROM N.A.
MEDLINE=81076554; PubMed=6777755;
Sikotav J., L., Auffray C., Rougeon F.;
"Structure of the constant and 3' untranslated regions of the murine
Balb/c gamma 2a heavy chain messenger RNA.";
Nucleic Acids Res. 8:3143-3155(1980).
[2]
SEQUENCE FROM N.A.
MEDLINE=81198976; PubMed=6262729;
Yamawaki-Kataoka Y., Miyata T., Honjo T.;
"The complete nucleotide sequence of mouse immunoglobulin gamma 2a gene
and evolution of heavy chain genes: further evidence for intervening
sequence-mediated domain transfer.";
Nucleic Acids Res. 9:1365-1381(1981).
[3]
SEQUENCE FROM N.A.
MEDLINE=81223894; PubMed=6787604;
Olio R., Auffray C., Mochamps C., Rougeon F.;
"Comparison of mouse immunoglobulin gamma 2a and gamma 2b chain genes
suggests that exons can be exchanged between genes in a multigenic
family.";
Proc. Natl. Acad. Sci. U.S.A. 78:2442-2446(1981).
[4]
MYELOMA PROTEIN MOPC 173.
MEDLINE=74175517; PubMed=4831970;
Bourgois A., Fougereau M., Rocca-Serra J.;
"Determination of the primary structure of a mouse IgG2a
immunoglobulin: amino-acid sequence of the Fc fragment. Implications
for the evolution of immunoglobulin structure and function.";
Eur. J. Biochem. 43:423-435(1974).
[5]
DISULFIDE BONDS.
MEDLINE=73056887; PubMed=4565406;
de Preval C., Fougereau M.;
"Determination of the primary structure of a mouse gamma G2a
immunoglobulin. Identification of the disulfide bridges.";
Eur. J. Biochem. 30:452-462(1972).

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; V00798; CAA24178.1; -;
PIR; A02152; G2MSA.
HSSP; P01842; 7FAB.
InterPro: IPR003006; Ig_MHC.
InterPro: IPR003597; Ig_cl.
InterPro: IPR003600; Ig_like.
Pfam: PF00047; Ig; 2.
SMART; SM00410; Ig_like; 1.
SMART; SM00407; IGC1; 2.
PROSITE; PS00290; IG_MHC; 1.
Immunoglobulin domain; Immunoglobulin C region.
NON_TER 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 144 204
FT DISULFID 250 308
FT MOD_RES 330 330 REMOVED POST-TRANSLATIONALLY.
SQ SEQUENCE 330 AA; 36389 MW; B84361C545A6864 CRC64;

Query Match 61.0%; Score 814; DB 1; Length 330;
Best Local Similarity 60.7%; Pred. No. 1.le-20;
Matches 150; Conservative 31; Mismatches 56; Indels 10; Gaps 2;

QY 10 VAEKLKFAFMDK-----THTCPP--CPAPELLGSGVFLFPKPKDKTLMISRTPEVT 59
Db 84 VAHPASSTKVDKKIEPRGPTIKPCPPCKCPAPNLLGGPSVFIFPPKIKDVLMSLSPIVT 143
QY 60 CVVVDVSHEDPEVKFNWYVDGVEVHNATKPREBOYNSTYRVVSVLTVLHODWLNQKEYK 119
Db 144 CVVVDVSEDDPDVOISWFWNNVEVHTAQOTQTHREDYINSTLRVVSALPIQIHDWMSGKEFK 203
QY 120 KYSNKALPAPIETKISKARGQPREPOVYTLPPSRDELTKNQSLTCLVKGFYPSDIAVE 179
Db 204 KYNKNDLPAPIETISKPGSVRAPOVYVLPPEPEEMTKQVTLTCMTDFMEDYIVE 263
QY 180 WESNGQENNYKTPPVLDSDGSPFLYSLKLTVDKSRWQGNVFCSVNHEALHNYTKQS 239
Db 264 WTNNKTELNYKTEPVLDSGDSYFMYSKLRVEKKNNVERNYSYSCSVVHGLHNHHTKS 323

QY 240 LSLSPGK 246
Db 324 FSRTPGK 330

RESULT 15
GCAB_MOUSE STANDARD; PRT; 335 AA.
ID GCAB_MOUSE AC P01864;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2A chain C region, B allele.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=82037861; PubMed=6170065;
RA Schreier P.H., Bothwell A.L.M., Mueller-Hill B., Baltimore D.;
RT "Multiple differences between the nucleic acid sequences of the
IgG2a and IgG2b alleles of the mouse.";
Proc. Natl. Acad. Sci. U.S.A. 78:4495-4499(1981).
RN [2]
RP SEQUENCE.
RX MEDLINE=82037777; PubMed=6794027;
RA Dognin M.J., Lauwereys M., Strosberg A.D.;
RT "Multiple amino acid substitutions between murine gamma 2a heavy
chain Fc regions of IgA and IgB allotypic forms.";
Proc. Natl. Acad. Sci. U.S.A. 78:4031-4035(1981).
CC -1- MISCELLANEOUS: THE SEQUENCE DIFFERS FROM THAT OF THE ALLELE,
FROM BALB/C MICE, AT 15% OF THE POSITIONS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; J00479; -; NOT_ANNOTATED_CDS.
PIR; A02153; G2MSAB.
HSSP; P01842; 7FAB.
InterPro: IPR003006; Ig_MHC.
InterPro: IPR003597; Ig_cl.
InterPro: IPR003600; Ig_like.
Pfam: PF00047; Ig; 3.
SMART; SM00410; IG_like; 1.
SMART; SM00407; IGC1; 2.
PROSITE; PS00290; IG_MHC; 1.
Immunoglobulin domain; Immunoglobulin C region.
NON_TER 1
FT SEQUENCE 335 AA; 36596 MW; FA3382792CBB13C6 CRC64;

Query Match		60.6%;	Score 809;	DB 1;	Length 335;
Best Local Similarity		55.8%;	Pred. No. 1.7e-20;		
Matches 144;		Conservative 39;	Mismatches 43;	Indels 32;	Gaps 1;
QY	21 KTHTC-----	-----PPCPAPELLGGPSVFLFPKPKD 48			
Db	78 OTITCNWAHPASSTKVYDKKIEPRVPI	TQNPCPPHQRVPPCAAAPDLLGGPSVFIFPKIKD 137			
QY	49 TLMISRTPEVTCVVVDVSHEDPEVKFNWYVDG	VEVHNAKTKPREQYNSTYRVVSVLT 108			
Db	138 VLMISLSPMWTVCVVVDSEDDVDQISW	FNVNVEVHTAQTHREDYNSTLRVVSALPIQ 197			
QY	109 HODWLNKEYCKVSNKALPAPIETISKAKG	QPREPQVYITLPPSRDELTKNQVSLTCLV 168			
Db	198 HODWMSGKEFKCKVNNRALPSPIETISK	PRGVRAPQVYVLPPEAEEMTKKEFSLTCTMI 257			
QY	169 KGFYPSDIAVENESNCOPENNYKTPPV	LSDSGSFYLSKLTVDKSRWQGNVFCSCVMH 228			
Db	258 TGFPAEIAVDWTSNGRTQNYKNTATVLD	SDGSYFMYSKLRVQKSTWERSLFCACSVVH 317			
QY	229 EALHNHYTKSLSLSPGK 246				
Db	318 EVLHNHLTTKTISRSLGK 335				

Search completed: April 21, 2003, 10:43:11
Job time : 10.4059 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: April 21, 2003, 10:42:21 ; Search time 36.1765 Seconds
(without alignments)
1401.120 Million cell updates/sec

Title: 7LINK2

Perfect score: 1334

Sequence: 1 DWLKAFYDKVAEKLKEAFMD.....MHEALHNHYTKSLSLSPCK 246

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1250	93.7	471	4	O8TC77
2	1242	93.1	701	4	Q96PQ8
3	1151	86.3	473	4	O8TC63
4	917	68.7	337	6	Q95M34
5	847	63.5	463	11	Q99LC4
6	843	63.2	469	11	O8R3V9
7	838	62.8	437	11	O9R1A4
8	829	62.1	473	11	O9B8L4
9	824	61.8	473	11	Q99L25
10	821	61.5	468	11	Q99L31
11	797	59.7	474	11	O8R3H6
12	794	59.5	473	11	O91Z05
13	486	36.4	34350	4	O8WZ42
14	463	34.7	26926	4	Q10466
15	463	34.7	26926	4	O8WZB3
16	406	30.4	16215	5	Q9NFS3

17	405	30.4	17352	5	Q95YM2
18	392	29.4	597	4	Q96BB9
19	390	29.2	7962	4	Q10465
20	387	29.0	597	4	Q9BQB8
21	387	29.0	597	4	Q9BU10
22	387	29.0	614	4	Q96GA6
23	379	28.4	375	4	Q9BSZ1
24	377	28.3	588	4	O8WUX4
25	377	28.3	613	4	O8WUK1
26	377	28.3	618	4	Q96AA6
27	375	28.1	613	4	Q96EY0
28	372	27.9	613	11	Q8VCX7
29	371	27.8	15281	3	Q09164
30	358	26.8	6632	5	O17361
31	357	26.8	6632	5	O17362
32	356	26.7	7107	5	Q9V4F7
33	352	26.4	5636	4	Q96RW7
34	347	26.0	5198	5	O76518
35	347	26.0	5604	4	O8WZ53
36	345	25.9	6831	5	O23550
37	345	25.9	7160	5	O23551
38	344	25.8	8563	2	Q54297
39	343	25.7	4824	5	Q95YM1
40	342	25.6	6658	5	O76281
41	342	25.6	6875	6	O28733
42	340	25.5	8805	5	O9V6V4
43	339	25.4	13055	5	Q09165
44	337	25.3	9376	2	O85168
45	331	24.8	7463	16	Q9Z4X6

ALIGNMENTS

RESULT 1

Q8TC77	ID	Q8TC77	PRELIMINARY;	PRT;	471 AA.
AC	O8TC77;				
DT	01-JUN-2002	(Tremblrel. 21, Created)			
DT	01-JUN-2002	(Tremblrel. 21, Last sequence update)			
DT	01-JUN-2002	(Tremblrel. 21, Last annotation update)			
DE	Hypothetical 51.8 kDa protein.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=SPLEEN;				
RA	Strausberg R.;				
RL	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; BC024289; AH24289.1;				
KW	Hypothetical protein.				
SQ	SEQUENCE 471 AA; 51791 MW; 388F7F4CF588660E CRC64;				
	Query Match	93.7%;	Score 1250;	DB 4;	Length 471;
	Best Local Similarity	57.0%;	Pred. No. 7.6e-44;		
	Matches 232;	Conservative	8;	Mismatches 6;	Indels 161; Gaps 2;
QY	1 DWLKA-----				5
	::: :				
Db	65 EWVSSMSSSSYYVADSVKGRFTISRDNAKNSLYLQMSLRARDTAVVYCARDLRQLTS 124				
QY	6 --FYD-----				8
	::: :				
Db	125 YWYFDLWGRGLTVTVSSASTKGPVSFFLPAPSSKTSGGTAALGCLVKDYFPEPVTVSWNS 184				
QY	9 -----				19
	-----KVAEKLKEAFM				
Db	185 GALTSGVHTTFAVLQSSGLYSLSSVTVVPSSSLGTQTYICNVNKKPSNTKVDKAVEPKSC 244				
QY	20 DKHTCTPCPAPELLGSPVFLFPKPKDITLMSRTPEVTCVVVDVSHEDPEVKFNWYVD 79				

```
Db 245 DKHTTCCPCAPPELLGGPSVFLFPKPKDTLMSRTPEVTCVVVDVSHEDPEVKFNWYVD 304
QY 80 GVEVHNAKTKPREQNSYRVSVLTLVHQLDNLNGKEYCKVSNKALPAPIETKTSKAK 139
Db 305 GVEVHNAKTKPREQNSYRVSVLTLVHQLDNLNGKEYCKVSNKALPAPIETKTSKAK 364
QY 140 GQREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 199
Db 365 GQREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 424
QY 200 DGSFFLYSKLTVDKSRWQGNFVSCSMHEALHNNHYTKQSLSPGK 246
Db 425 DGSFFLYSKLTVDKSRWQGNFVSCSMHEALHNNHYTKQSLSPGK 471

RESULT 2
Q96P08
ID Q96P08 PRELIMINARY: PRT: 701 AA.
AC Q96P08;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-WAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Factor VII active site mutant immunocoujugate.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21477448; PubMed=11593034;
RA Hu Z., Garen A.;
RT "Targeting tissue factor on tumor vascular endothelial cells and tumor
RT cells for immunotherapy in mouse models of prostatic cancer.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12180-12185(2001).
DR EMBL: AF272774; AAKS686.1; -.
DR InterPro: IPR000152; ASX_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF 2.
DR InterPro: IPR001881; EGF Ca.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001254; Serp_protease_Try.
DR InterPro: IPR000294; Vitk_dep_GLA.
DR Pfam: PF00008; EGF; 2.
DR Pfam: PF00594; gla; 1.
DR Pfam: PF00047; ig; 2.
DR Pfam: PF00089; trypsin; 1.
DR SMART: SM00181; EGF; 2.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF 2; UNKNOWN_1.
DR PROSITE: PS01187; EGF Ca; UNKNOWN_1.
DR PROSITE: PS00011; GLU CARBOXYLATION; UNKNOWN_1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 701 AA; 77826 MW; 94AC6CEB42CC992F CRC64;

Query Match 93.1%; Score 1242; DB 4; Length 701;
Best Local Similarity 87.2%; Pred. No. 6.5e-43;
Matches 231; Conservative 3; Mismatches 4; Indels 27; Gaps 2;

QY 1 DMLKAFYDKVAEKLKEA-----FMDKTHTCPCAPPELLGGPSVFL 41
Db 445 EWL-----QKLMRSEPRGVLLRAPFGSAEPKSCDKTHTCPCAPPELLGGPSVFL 496
QY 42 FPPKPKDTLMSRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRV 101
Db 497 FPPKPKDTLMSRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRV 556
QY 102 VSVLTVLHODWLNKREYCKVSNKALPAPIETKTSKAKQPREPVYTLPPSRDELTKNQ 161
```

```
Db 557 VSVLTVLHODWLNKREYCKVSNKALPAPIETKTSKAKQPREPVYTLPPSRDELTKNQ 616
QY 162 VSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFPLYSKLTVDKSRWQGNV 221
Db 617 VSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFPLYSKLTVDKSRWQGNV 676
QY 222 FSCSMHEALHNNHYTKQSLSPGK 246
Db 677 FSCSMHEALHNNHYTKQSLSPGK 701

RESULT 3
Q8TC63
ID Q8TC63 PRELIMINARY: PRT: 473 AA.
AC Q8TC63;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 52.0 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
KW EMBL; BC025985; AAH25985.1; -.
KW Hypothetical protein.
SQ SEQUENCE 473 AA; 51986 MW; E29920B09BA369F5 CRC64;

Query Match 86.3%; Score 1151; DB 4; Length 473;
Best Local Similarity 54.2%; Pred. No. 1.2e-39;
Matches 217; Conservative 16; Mismatches 13; Indels 154; Gaps 4;

QY 1 DML-----KAFYDKVAB----- 12
Db 74 EWIGTINFGSNMYSPSLRSRVMTSADMSKNSFYLKLDSTVTAADTAVYYCAAGHLVMGFG 133
QY 13 -----KLUKAFMD----- 20
Db 134 AHWGQGLVSVSPASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPTVYSWNSGALT 193
QY 21 -----KTHF----- 26
Db 194 SCVHTFPFVAVLQSSGLYSLSVVTVFSSSLGTQTYTCNVDDHKPSNTKVDKRVESKYGPPCP 253
QY 27 PCPAPPELLGGPSVFLFPPKPKDTLMSRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNA 86
Db 254 SCPAPEELGGPSVFLFPPKPKDTLMSRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNA 313
QY 87 KTKPREEQYNSTYRVVSVLTVHLQDWLNKGEYCKVSNKALPAPIETKTSKAKQPREPQ 146
Db 314 KTKPREEQFNSTYRVVSVLTVHLQDWLNKGEYCKVSNKGLPSSIEKTIKTSKAKQPREPQ 373
QY 147 VYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFELY 206
Db 374 VYTLPPSPEEEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFELY 433
QY 207 SKLTVDKSRWQGNFVSCSMHEALHNNHYTKQSLSPGK 246
Db 434 SRLTVDKSRWQGNFVSCSMHEALHNNHYTKQSLSPGK 473

RESULT 4
Q9SM34
ID Q9SM34 PRELIMINARY: PRT: 337 AA.
AC Q9SM34;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Immunoglobulin gamma 1 heavy chain constant region
```



```

DE (Fragment).
GN IGH1.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
DR [1]
DR PROSITE: PS003006; Iq_MHC.
RP SEQUENCE FROM N.A.
RA Wagner B.; (DEC-2000) to the EMBL/GenBank/DBJ databases.
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98383416; PubMed=9717671;
RA Wagner B.; Overesch G.; Sheoran A.; Holmes M.; Richards C.;
RA Leibold W.; Radbruch A.;
RT "organization of the equine immunoglobulin heavy chain constant region
RT genes. III. Alignment of c-mu, c-gamma, c-epsilon and c-alpha genes.";
RL Immunobiology 199;105-119(1998).
DR EMBL; AJ300675; CAC44624.1;
DR InterPro; IPR003006; Iq_MHC.
DR Pfam; PF00047; Iq; 2.
DR PROSITE; PS00290; Iq_MHC; UNKNOWN_2.
FT NON_TER
SQ SEQUENCE 337 AA; 37438 MW; A60BF2B01DEFD1F6 CRC64;

Query Match 68.7%; Score 917; DB 6; Length 337;
Best Local Similarity 63.8%; Pred. No. 3.4e-30;
Matches 169; Conservative 33; Mismatches 40; Indels 23; Gaps 4;

QY 2 WLKAFY-----DKVAEKL-----KEAFDMKTHTCPCPAPELLGGPSVFLFP 43
DB 76 WTSEYICNVVHAASNEKFKVRIPIPDNHQKVCDSK---CPKCPAPELGGPSVIFP 132
QY 44 PKPKDTLMIKTRPTVCTVVDVSHEDPEVKFNWYDGVGVHNAKTKPREQYNSTYRVVS 103
DB 133 PNPKDTLMIKTRPTVCTVVDVSHEDPEVKFNWYDGVGVHNAKTKPREQYNSTYRVVS 192
QY 104 VLTVLHODWLNGLGKEYCKVSNKALPAPIETKISAKQPREQVYTLPPSRDELTKNQVS 163
DB 193 VLRIOHODWLNGLGKEYCKVSNKALPAPIETKISAKQPREQVYTLPPSRDELTKNQVS 252
QY 164 LTCVLVGFYSDIAVESNGQP--ENNYKTPPVLDSDGFFLYSKLTVDKSRWQGNV 221
DB 253 VTLVLDGFFLYSKLTVDKSRWQGNV--ENNYKTPPVLDSDGFFLYSKLTVDKSRWQGNV 312
QY 222 FSCSVMEALHNHYTQKSLSLSPCK 246
DB 313 FTGCVMEALHNHYTQKSLSLSPCK 337

RESULT 5
Q99LC4 PRELIMINARY; PRT; 463 AA.
AC Q99LC4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to RIKEN CDNA 181006O009 gene.
GN IGH-4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
DR [1]
DR PROSITE: PS003006; Iq_MHC.
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN EMBL; BC003435; AAH03435.1;
DR HSSP; P01842; 7FAB.
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR003599; Iq.
DR InterPro; IPR003597; Iq-cl.
DR InterPro; IPR003600; Iq_like.

```

```

DR InterPro; IPR003006; Iq_MHC.
DR InterPro; IPR003596; Iq_v.
DR Pfam; PF00047; Iq; 4.
DR SMART; SM00409; Iq; 2.
DR SMART; SM00407; Iq; 2.
DR SMART; SM00406; Iq; 1.
DR SMART; SM00410; Iq; 1.
DR PROSITE; PS00290; Iq_MHC; UNKNOWN_1.
SQ SEQUENCE 463 AA; 51007 MW; EAA674C6BBC30783 CRC64;

Query Match 63.5%; Score 847; DB 11; Length 463;
Best Local Similarity 37.1%; Pred. No. 9.5e-27;
Matches 154; Conservative 46; Mismatches 39; Indels 176; Gaps 8;

QY 2 WLK-----AEYDKVAEKLK-----EAFMD----- 20
DB 55 WKQRTQGLEWGEIYVPGSGNTYY---SEKFKGKATLTDDKSSSTAYMHLSSLTSEDSA 111
QY 21 -----KT----- 22
DB 112 VYFCARSSYSYDLFAYWGGTGLTVSAAKTPPSVYPLAPGSAQAQTNMWTGLGLVKGY 171
QY 23 -----HT----- 24
DB 172 FPEPVTVMNSGSLSSGVHTFFPAVLQSDLYTLSSSVTPVSPSTWTPSETVTCNVAHPASSTK 231
QY 25 -----CPD--CPAPELGGPSVFLPPKPKDTLMSRTPETVTVVDVSHEDPE 71
DB 232 VDKKIVPRDGCGRPCICTVPEV---SSVFIKPKPKDVLTITLTPKVTVCVVDVTSKDDPE 288
QY 72 VKENWYVDGVEVHNAKTKPREQYNSTYRVVSVLTVLHODWLNGLGKEYCKVSNKALPAPI 131
DB 289 VQFSWEVDVVEVHTAQTQPREQYNSFRSVSELPIMHODWLNGLGKEYCKVSNKALPAPI 348
QY 132 EKTISKAGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVESNGQYENNYK 191
DB 349 EKTISKAGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVESNGQYENNYK 408
QY 192 TTPVLDSDGFFLYSKLTVDKSRWQGNV--ENNYKTPPVLDSDGFFLYSKLTVDKSRWQGNV 246
DB 409 NTQPIMDTGSYFIYSLNVOKSNWEAGNFTCSVLHGLHNHHTKSLSHSPCK 463

RESULT 6
Q8R3V9 PRELIMINARY; PRT; 469 AA.
AC Q8R3V9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 52.0 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
DR [1]
DR PROSITE: PS00290; Iq_MHC.
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN EMBL; BC024405; AAH24405.1;
DR EMBL; BC024405; AAH24405.1;
KW Hypothetical protein.
SQ SEQUENCE 469 AA; 51976 MW; 534793F155D05457 CRC64;

Query Match 63.2%; Score 843; DB 11; Length 469;
Best Local Similarity 36.8%; Pred. No. 1.5e-26;
Matches 150; Conservative 46; Mismatches 47; Indels 165; Gaps 5;

QY 1 DWL----- 3
DB 65 EWJGFIKANGYTYEYSASVKGRFTISRDNQSQSLYLQMNALRAEDSATYICARDRSS 124
QY 4 -----KAFYDK----- 9

```

```

Db 125 YYTSGTFAFWGOGTLVTVSAAKTTPPSVYPLAPGSAQTNSMVTGLCLVKGYPFPTV 184
QY 10 -----VAEKLKEAFMDKTHT - 24
Db 185 TWSGSLSSGVHTFPVAVLQSLDLYTLSSSVTVPSSTWPSVTCNVNHAHPASSTKVDKKIYP 244
QY 25 -----CPP--CPAPELLGGPSVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWTV 78
Db 245 RDCGCKPCTCTPEV---SSVFIFPPKPKDVLITITLTPKVTCCVVVDISKDDPEVQFSWFV 301
QY 79 DGYEVINAKTKPREOYNSTYRVSVLTVLHQDLNGKEYCKVSKNALPAPIEKTISKA 138
Db 302 DDYEVHTAQTKPREOYNSTYRVSVLTVLHQDLNGKEYCKVSKNALPAPIEKTISKA 361
QY 139 KGOPREPOVYTLPPSRDELTKNOVSLTCLVKGYPSPDIKAVESNGOPENNYKTPPVLD 198
Db 362 KGRKAPQVITPPPREQWAKQKVSCTMTDFFPEDITVWQNGQPAENYKNTQPIMD 421
QY 199 SDGSFFLYSKLTVKSRWQOQNVFSCVMHEALHNYTKQSLSLSPGK 246
Db 422 TDGSYFVSKLVNOKSNWEAGNTFTCSVLHGLHNTKSLSHSPGK 469

RESULT 7
Q9R1A4
ID Q9R1A4 PRELIMINARY; PRT; 437 AA.
AC Q9R1A4
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Gamma1 heavy chain of Mab7 (Fragment).
GN IGH-4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilde K.G., Yu X., Ekramoddoullah A.K.M., Misra S.;
RT "Cloning of cDNAs encoding for anti-white pine blister rust monoclonal
RT antibody (Mab 7, its light and heavy chains) and construction of a
RT single chain antibody (scFv).";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF152372; AAD40243.1; -.
DR HSSP; P01842; 7FAB.
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR003600; Igh-like.
DR InterPro; IPR003006; Igh_MHC.
DR Pfam; PF00047; Igh-4.
DR SMART; SM00406; Igh-4.
DR SMART; SM00410; Igh-like; 2.
DR PROSITE; PS00290; IGH_MHC; UNKNOWN_1.
FT NON_TER
FT NON_TER
SQ SEQUENCE 437 AA; 48142 MW; 5C3A7BB3EE7D697C CRC64;

Query Match 62.8%; Score 838; DB 11; Length 437;
Best Local Similarity 37.4%; Pred. No. 1.9e-26;
Matches 148; Conservative 48; Mismatches 47; Indels 153; Gaps 5;

QY 1 DWLKAF-----YDKVAEKLKEAFMDK----- 21
Db 45 ENWASFSGGIYITDSVKGRTTIYKDKDRNLTSLQMSSLSRSEDYAMVYCARDYSAYWG 104
QY 22 -----TH 23
Db 105 PGLTVTVSAAKTTPPSVYPLAPGSAQTNSMVTGLCLVKGYPFPTVTVNWSGSLSGVH 164
QY 24 T-----CPP--CPA 30
Db 165 TFPVAVLQSLDLYTLSSSVTVPSSTWPSVTCNVNHAHPASSTKVDKKIYPRDCCCKPCICTV 224

```

```

QY 31 PELLGGPSVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKP 90
Db 225 PEV---SSVFIFPPKPKDVLITITLTPKVTCCVVVDISKDDPEVQFSWFVDDVEVHTAQTQP 281
QY 91 REEQNSTYRVSVLTVLHQDLNGKEYCKVSKNALPAPIEKTISKAKGQPREPOVYTL 150
Db 282 REEQNSTFRSVSELPIMHQDLNGKEYCKVSKNALPAPIEKTISKAKGQPREPOVYTL 341
QY 151 PPSRDELTKNOVSLTCLVKGYPSPDIKAVESNGOPENNYKTPPVLDSDGSFFLYSKLTP 210
Db 342 PPKKQAKMDKVSCTMTDFFPEDITVWQNGQPAENYKNTQPIMDTDGSFFLYSKLTP 401
QY 211 VDKSRWQOQNVFSCVMHEALHNYTKQSLSLSPGK 246
Db 402 VQKSNWEAGNTFTCSVLHGLHNTKSLSHSPGK 437

RESULT 8
Q9D8L4
ID Q9D8L4 PRELIMINARY; PRT; 473 AA.
AC Q9D8L4
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE 181006009Rik protein.
GN IGH-1 OR 181006009Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=PANCREAS;
RA MEDLIN=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gliss C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staebli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz T., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK007918; BAB25349.1; -.
DR HSSP; P01842; 7FAB.
DR MGD; MGI:96443; Igh-1.
DR InterPro; IPR003599; Igh-1.
DR InterPro; IPR003597; Igh-cl.
DR InterPro; IPR003600; Igh-like.
DR InterPro; IPR003006; Igh_MHC.
DR Pfam; PF00047; Igh-4.
DR Pfam; PF00047; Igh-4.
DR SMART; SM00409; Igh-4.
DR SMART; SM00407; Igh-4.
DR SMART; SM00406; Igh-1.
DR SMART; SM00410; Igh-like; 1.
DR PROSITE; PS00290; IGH_MHC; UNKNOWN_1.
SQ SEQUENCE 473 AA; 51699 MW; 9DED57A514475FBB CRC64;

Query Match 62.1%; Score 829; DB 11; Length 473;
Best Local Similarity 36.4%; Pred. No. 5.9e-26;
Matches 154; Conservative 43; Mismatches 46; Indels 180; Gaps 6;

```

```
QY 1 DWLK-----AFYDKVAEKLK-----EAFMD----- 20
Db 54 NWKQRPQGLEWIGKIGPGSGSTYYN---EKFKGATLTADKSSSTAYMQLSLTSEDS 110
QY 21 -----KT----- 22
Db 111 AVYFCARSGYDYDFAYWQGTTLVTVSAAKTTAPSVVPLAPVCGGTGSSVTGLCLVKGY 170
QY 23 -----HT----- 24
Db 171 FPEPVTLTWNSGLSSGVTHTFPALLOGLYTLSSSVTVTSNTWPSQITITCNVAHPASSTK 230
QY 25 -----CPPCAPPELLGGPSVFLFPKPKDITLMISRTPEVTCVVV 63
Db 231 VDKKIEPRVITQPCPPLEKPCPCAPDLGGPSVFIKPKIKDVLMLISLSPMVTVCVV 290
QY 64 DVSHEDPEVFNKVVGVVHNAKTRPREQYNSTRVSVLVHLDWLNKGEYKCKVS 123
Db 291 DVSEDDPDVQISWVFNNEVHTAQOTQTHREDYNSTLRVVSALPIQHODMNSGKEFKCKVN 350
QY 124 NKALPAPIETISKAKQPREPOVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVEWESN 183
Db 351 NRALPSPIETISKPRGPVAPQYVILPPPAEEMTKKQVTLTCHMTGFLPAEIAVDWTSN 410
QY 184 GOPENNYKTTPPVLDSDGSFLLXSKLTVDKSRWQOQNVFSCSVHMEALHNHYTKQSLSL 243
Db 411 GRTEQNYKNTATVLDSDGSYFMSKLRVQKSTWERSLFCACSVVHGLHNLTKTISR 470
QY 244 PGK 246
Db 471 LGK 473
```

```
RESULT 9
Q99L25 PRELIMINARY; PRT; 473 AA.
AC Q99L25;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to RIKEN CDNA 1810060009 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003888; AA03888.1; -
DR HSSP; P01842; 7FAB.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003600; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00409; IGV; 1.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG-like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 473 AA; 52449 MW; BE9889B7986DA155 CRC64;

Query Match 61.8%; Score 824; DB 11; Length 473;
Best Local Similarity 38.6%; Pred. No. 9.7e-26;
Matches 158; Conservative 33; Mismatches 55; Indels 163; Gaps 6;

QY 1 DWLKAFY-----DKVAEKLK-----EAFM----- 19
Db 65 EWIGIYIPRGSTKYNEKFKGATLTADKSSSTAYMQLNSLTSEDSAVCFCSRGSIYIG 124
```

```
QY 20 -----DKT----- 22
Db 125 YGLYFDYWGQTTITVSSAKTTAPSVVPLAPVCGDTTGGSSVTGLCLVKGYFPEPVTLTW 184
QY 23 -----HT----- 24
Db 185 NSGSLSSGVHTFPVAVLQSDLYTLSSSVTVTSSTWPSQSITCNVAHPASSTKVKDKIEPRG 244
QY 25 -----CPP-----CAPPELLGGPSVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNW 77
Db 245 PTKPCPCPCAPNLLGGPSVFIKPKIKDVLMLISLSPMVTVCVVVDVSEDDPDVQISWF 304
QY 78 VDGVEVHNAKTRPREQYNSTRVSVLVHLDWLNKGEYKCKVSNKALPAPIETISK 137
Db 305 VNNVEVLTAQTQTHREDYNSTLRVVSALPIQHODMNSGKEFKCKVNNKALPAPIETISK 364
QY 138 AKQPREPOVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVEWESNQCPNNYKTTTPVL 197
Db 365 PKGSVRAPQYVILPPPAEEMTKKQVTLTCHMTGFLPAEIAVDWTSN 424
QY 198 DSDGSFLXSKLTVDKSRWQOQNVFSCSVHMEALHNHYTKQSLSLSPGK 246
Db 425 DSDGSYFMSKLRVEKKNWVERNSYSCSVVHGLHNLTKTSRTEPK 473

RESULT 10
Q99L31 PRELIMINARY; PRT; 468 AA.
AC Q99L31;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to RIKEN CDNA 1810060009 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003878; AA03878.1; -
DR HSSP; P01842; 7FAB.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003600; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00409; IGV; 1.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG-like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 468 AA; 51661 MW; 96352328B3332ADB CRC64;

Query Match 61.5%; Score 821; DB 11; Length 468;
Best Local Similarity 37.2%; Pred. No. 1.2e-25;
Matches 154; Conservative 33; Mismatches 58; Indels 169; Gaps 4;

QY 2 WLK-----AFYD----- 4
Db 55 WVKQRPQGLEWIGWIDPDEGETKYAPFQDKATITADTSSNTAYLQSLTSEDYIYY 114
QY 5 -----AFYD----- 8
Db 115 CARNLLYGGYDYWGQGTITVSSAKTTAPSVVPLAPVCGDTTGGSSVTGLCLVKGYFPEP 174
QY 9 -----KVAEKLKFAFMK- 21
Db 175 VTLTNWNSGLSSGVHTFPVAVLQSDLYTLSSSVTVTSSTWPSQSITCNVAHPASSTKVDK 234
QY 22 -----THTCPP--CPAPPELLGGPSVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEV 72
```

```

235 IEPRGTIRKPCPKCPAPNLLGGPSVFIFFPKIKDVLMSLSPMVTCVVVDVSEDDPDV 294
QY 73 KFNWYDGVVEVHNKTKPREEQNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIE 132
DB 295 QISWFVNNVEVLTQAQTQTHREDYNSTLRVVSALPIQHODWMSGKEFKCKVNNKALPAPIE 354
QY 133 KTISKAKGQPREPQVYITLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKT 192
DB 355 RTISKPKGSVRAPQVYIILPPPEEMTKQVTLTCMTDWPEDYIYVETWNNKGTENLYKN 414
QY 193 TTPVLDSGSGFFLYSLKTLVDKSRMOQGNVFCSCVMHEALHNYTKQKLSLSPGK 246
DB 415 TEPVLDSGSGFYMSKLRVEKKNWERNYSVCSVVEHGLNHHHTTKFSRTPGK 468

RESULT 11
Q8R3H6 PRELIMINARY; PRT; 474 AA.
AC Q8R3H6;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical 51.7 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025447; AAH25447.1; -.
KW Hypothetical protein.
SQ SEQUENCE 474 AA; 51748 MW; 8608B57C6CD2874A CRC64;
```

```

Query Match 59.7%; Score 797; DB 11; Length 474;
Best Local Similarity 35.2%; Pred. NO. 1.4e-24;
Matches 148; Conservative 40; Mismatches 58; Indels 175; Gaps 4;

QY 1 DWLKA-----FYD----- 5
DB 54 NWKRRPGKLEWIGRIFPGDGTTHYSGKFGKAKLTADKSSVTAPLQLTSLTSEDSAVY 113
QY 6 -----FYD----- 8
DB 114 FCARDSYGDYFDDMGQGATVTVSSAKTTPPSVYPLAPCGDGTGSSVTLGCLVKGYTPE 173
QY 9 -----KVAEKLKFAFMDK 21
DB 174 SVTVTNWNSGSLSSVHTFPALLQSLGYTMSSSVTPSPSTQVTCVSAHPASSTTVDK 233
QY 22 -----THTCPP-----CPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDV 65
DB 234 KLEPSPISTINPCPCCKECHKCAPNLEGGPSVFIFPPNKKDLMLSLTPKVTCTVVVDV 293
QY 66 SHEDPEVFNWYDGVVEVHNKTKPREEQNSTYRVSVLTVLHODWLNKGEYKCKVSNK 125
DB 294 SEDDDPDVQISWFVNNVEVLTQAQTQTHREDYNSTLRVVSALPIQHODWMSGKEFKCKVNNK 353
QY 126 ALPAPIETISKAKGQPREPQVYITLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQ 185
DB 354 DLPSPFIETISKIKGLVRAPQVYIILPPPAEQLSRKDVSLTCLVGFNPGDISVETWSNGH 413
QY 186 PENNYKTTTPVLDSGSGFFLYSLKTLVDKSRMOQGNVFCSCVMHEALHNYTKQKLSLSPG 245
DB 414 TEENYKADTAPVLDSGSGFYIYKSLDKITKWEKTSFSCNVRHEGLKNLYLKTISRSPG 473
QY 246 K 246
DB 474 K 474
```

```

RESULT 12
Q91Z05 PRELIMINARY; PRT; 473 AA.
AC Q91Z05;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical 51.9 kDa protein.
GN AU044919.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010327; AAH10327.1; -.
DR MGI; MGI:2144967; AU044919.
DR InterPro; IPR000345; CytC_heme_bind.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_3.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 473 AA; 51946 MW; CF625F008932AF12 CRC64;
```

```

Query Match 59.5%; Score 794; DB 11; Length 473;
Best Local Similarity 41.8%; Pred. NO. 1.8e-24;
Matches 148; Conservative 39; Mismatches 58; Indels 109; Gaps 4;

QY 2 WL-----KAFYDK----- 9
DB 120 WLRRIDYWGQGTITVSSAKTTPPSVYPLAPCGDGTGSSVTLGCLVKGYPSVTVTN 179
QY 10 -----VAEKLKFAFMDK----- 21
DB 180 SGLSSSVHTFPALLQSLGYTMSSSVTPSPSTQVTCVSAHPASSTTVDKLEPSGP 239
QY 22 -----THTCPP-----CPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEV 72
DB 240 ISTINPCPCCKECHKCAPNLEGGPSVFIFPPNKKDLMLSLTPKVTCTVVVDVSEDDPDV 299
QY 73 KFNWYDGVVEVHNKTKPREEQNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIE 132
DB 300 QISWFVNNVEVLTQAQTQTHREDYNSTLRVVSALPIQHODWMSGKEFKCKVNNKALPAPIE 359
QY 133 KTISKAKGQPREPQVYITLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKT 192
DB 360 RTISKIKGLVRAPQVYIILPPPAEQLSRKDVSLTCLVGFNPGDISVETWSNGHTENYKD 419
QY 193 TTPVLDSGSGFFLYSLKTLVDKSRMOQGNVFCSCVMHEALHNYTKQKLSLSPGK 246
DB 420 TAPVLDSGSGFYIYKSLDKITKWEKTSFSCNVRHEGLKNLYLKTISRSPGK 473
```

```

RESULT 13
Q8WZ42 PRELIMINARY; PRT; 34350 AA.
AC Q8WZ42;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Ttin.
GN TTN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Freiburg A., Trombitas K., Hell W., Cazorla O., Fougereousse F.,
```

RA Centner T., Kolmerer B., Witt C., Beckmann J.S., Gregorio C.C.,
RA Granzier H., Labeit S.;
RT "Series of exon-skipping events in the elastic spring region of titin
RL as the structural basis for myofibrillar elastic diversity.";
RN Circ. Res. 86:1114-1121(2000).
RP [2]
RX SEQUENCE FROM N.A.
RX MEDLINE=21573839; PubMed=11717165;
RA Bang M.L., Centner T., Fornoff F., Geach A.J., Gotthardt M.,
RA McNabb M., Witt C.C., Labeit D., Gregorio C.C., Granzier H.,
RA Labeit S.;
RT "The complete gene sequence of titin, expression of an unusual ~700
RT kDa titin isoform and its interaction with obscurin identify a novel
RT z-line to I-band linking system.";
RL Circ. Res. 89:1065-1072(2001).
DR EMBL: AJ277892; CAD12456.1; .
DR InterPro: IPR000282; Cytok_receptor_2.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR000577; FGGY_kin.
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR001092; HLH_basic.
DR InterPro: IPR003599; Ig_c2.
DR InterPro: IPR003598; Ig_c2.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR InterPro: IPR002016; Peroxidase.
DR InterPro: IPR004168; PPAK_motif.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00041; fn3; 132.
DR Pfam: PF00047; Ig; 146.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF02818; PPAK; 53.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00060; FN3; 133.
DR SMART: SM00409; IG; 167.
DR SMART: SM00408; IG2; 148.
DR SMART: SM00406; IGV; 23.
DR SMART: SM00220; S_TKC; 1.
DR SMART: SM00219; TYR_KG; 1.
DR PROSITE: PS00933; FGGY_KINASES_1; UNKNOWN_1.
DR PROSITE: PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
DR PROSITE: PS00435; PEROXIDASE_1; UNKNOWN_1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
SQ SEQUENCE 34350 AA; 3816262 MW; 5B1120058A7CE58A CRC64;

Query Match 36.4%; Score 486; DB 4; Length 34350;
Best Local Similarity 0.5%; Pred. No. 5,1e-05;
Matches 164; Conservative 37; Mismatches 44; Indels 30646; Gaps 47;
Qy 2 WLK----- 4
|||
Db 3001 WLKNGVEIKTDKQMRTKLTHSLNIRNVHFGDAADYTFVAGKATSTATLYVEARHIEF 3060
Qy 5 ----- 4
Db 3061 RKHKIDIKVLEKKRAMFEVSEPDITVQWMKDDQELQITDRIKIQEKYVHRLIPSTR 3120
Qy 5 ----- 4
Db 3121 MSDAGKYTVAGNVSTAKLFEVGRDVRISIKKEVQVIEKQRAVVEFEVNEDDVDHAWY 3180
Qy 5 ----- 4
Db 3181 KDGIEINQVOERHKYVVERRIHRMFISETRQSDAGEYTFVAGNRSSVTLVNAPEPPQ 3240
Qy 5 ----- 4
Db 3241 VLQELQPTVVGSGKPARFCAVIGSRPQPKISWYKEEQLLSTGFKCKFLHDGQETLLIE 3300
Qy 5 ----- 4

Db 3301 AFPDAAVYTCEAKNDYGVATTSSASLSVEVPEVSWDQEMPVYPPIAITPLQDVTSEGG 3360
Qy 5 ----- 4
Db 3361 PARFQCRVSGTDLKVSWSYKDKKIKPSRFRMTOPEDTYQLEIAEAYPEDEGTYTFVASN 3420
Qy 5 ----- 4
Db 3421 AVGOVSSPANLSLEAPESILHERIEQETEMEMKEFSSFLSAEERGLHSAEQLQSKINET 3480
Qy 5 ----- 4
Db 3481 LELLSESPYSTKFDSEKGTGPIFKEVSNADISMGDVATLSVTVIGIPKPIQWFFNG 3540
Qy 5 ----- 4
Db 3541 VLLTPSADYKVFVGDGDDHSLIILFTKLEDEGEYTCMASNDYKGTICSAYLKINSKGEGHK 3600
Qy 5 ----- 4
Db 3601 DTETESAVAKSLEKLGPCPPHFLKELKPIRCAOGLPAIFETVVVGEPAITVTFWKENKQ 3660
Qy 5 ----- 4
Db 3661 LCTSVYTYTIHNPNGSGTFIVNDPOREDSGLYICKAENMLGESTCAEALLVLEDTMTD 3720
Qy 5 ----- 4
Db 3721 TPKAKSTPEAPDFPOTPLKGPVAVELDSQEIATFVKDTILKAALITEENQQLSYEHI 3780
Qy 5 ----- 4
Db 3781 AKANELSSQLPLGAQELQSLIQLDQKLTPESTREFLCINGSIHQPLKEPSNMLQIIVOS 3840
Qy 5 ----- 4
Db 3841 QKTSKEGILPEPETOAVLSDTEKIFPSAMSIQINSLTVEPLKTLAEPEGNYPOSS 3900
Qy 5 ----- 4
Db 3901 IEPPHMSYLTVAEEVLSPKKTVDNTNREQRTVLQKQASALISLQSLAEGHVESLQS 3960
Qy 5 ----- 4
Db 3961 PDVMSQVNYEPLVPSEHSCTEGGKILLIESANPLENAGODSAVRIEKGSLRFFPLALEEK 4020
Qy 5 ----- 4
Db 4021 QVLLKEHSDNVWMPDQIIESKREPVAIKKQVOYGRDLLSKESLLSGIPEEQRLNLKI 4080
Qy 5 ----- 4
Db 4081 QICRALQAASAEQPLFSEWLRIEKEVAVNTQEPHIMCMYLTSAKSVTEVTI 4140
Qy 5 ----- 4
Db 4141 IIEDVDPOMANKMELRDALCAIIEEDILTAEGPRIQQAGKATSLQEMDSFSGSQKVE 4200
Qy 5 ----- 4
Db 4201 PITEPEVESKYLISPEEVSYFNVQSRVKYLDATPVTKGVAVWVDEKQDESLKPSBEKE 4260
Qy 5 ----- 4
Db 4261 ESSSESGTEEVATVKIQEAEGGFIKEDGPMIHTPLVDVTVEEGDIVHLTTSTTNAKEYNW 4320
Qy 5 ----- 4
Db 4321 YFENKLVPSDEKFKCLQDQNTYTLVIDKVNTEHDGQEVCEALNDSGKATSAKLTVVVKR 4380
Qy 5 ----- 4

Db 4381 AAPVKKRIBPLEVALGHLAKFTCEIQSAPNVRQWFKAGREIYESDKCSIRSKYISSL 4440
QY 5 ----- 4
Db 4441 EILRTQVDCGEYCKASNEYGSVCTATLTVTEAYPPFLSRPKSLTFVGAKEFICT 4500
QY 5 ----- 4
Db 4501 VTGTPVETIWOQDGAALSPNWKISDAENKHILELSNLTIQDRGVYSKASNKFGADI 4560
QY 5 -----A 5
Db 4561 CQBELIIDKPHFIKELEPVQSAINKVKVHLECOVDEDRKVTVTWSDGQKLPFGDYKIC 4620
QY 6 FYDKVA----- 11
| | | |
Db 4621 FEDKIATLEIPLAKLDSGTIVCTASNEAGSSCSATVTVREPPSFVKVDPYSILMPGE 4680
QY 12 ----- 11
Db 4681 SARLHCKLKGSPVIQVTFWKNKSELSESTVRMYFNSEAILDITDVKVEDSGYSCEAV 4740
QY 12 ----- 11
Db 4741 NDVGSDSCSTEIVIKPEPPFIKLEPADIVRGTNALLQCEVSGTGPFISWFKDKKOIRS 4800
QY 12 ----- 11
Db 4801 SKYRLFSQKSLVCLFEIFSADVGEYECVWANEVKGCGMATHLLKEPPTVKVDDL 4860
QY 12 ----- 11
Db 4861 IALGGOTVTLQAAVRGSEPTVTWKGQEVIREDKIKMSFNGVAVLIIPDVQISFGGK 4920
QY 12 ----- 11
Db 4921 YTCIAENAGSQTSGELIVKEPAKIIERAEIQTAGDPATLEYTVAGTPPELKPWKYKD 4980
QY 12 ----- 11
Db 4981 GRPLVASKKYRISFKNNVAQLKFYSAEHLDSGQYTFEISNEVGSSCETFTFVLDRDIAP 5040
QY 12 ----- 11
Db 5041 FFTKPLRNVDSVVGTCRLDCKIAGSLPMRVSMFKDGEIAASDRYIAFVEGTASLEII 5100
QY 12 ----- 11
Db 5101 RVDMDAGNFTCRATNSVSGSKDSSGALIVOEPPSFVTKPGSKDVLPGSAVCLKSTFGST 5160
QY 12 ----- 11
Db 5161 PLTIRWFGKNKELVSGSCYITKEALESLELYLVKTSQYTYCKVSNVAGGVECSANL 5220
QY 12 ----- 11
Db 5221 FVKEPATFVEKLEPSQLLKKGDATQACKVTGTPPIKITWFANDREIKESKHRMSFVES 5280
QY 12 ----- 11
Db 5281 TAVLRITDVGIEDSGEYMC EAONEAGSDHCSSIVIVKESPYTFKEPKPIEVLKEYDVMILL 5340
QY 12 ----- 11
Db 5341 AEVAGTPPFEITWFKDNTILRSRKYKTFIQDHLVLSQLILKFVAADAGEYQCRVTNEVGS 5400
QY 12 ----- 11
Db 5401 SIGSARVTLREPPSFIIKKEIESTSLRGTAFAQATLKGSLPITVTWTKDSDEITEDDNIR 5460
QY 12 ----- 11
Db 5461 MTFENNVASLYLGGIEVKHDKGVYCOAKNDAGIQRCSALLSVKEPATITEAVSIDVTG 5520

QY 12 ----- 11
Db 5521 DPATLOVKFSGTKEITAKWFKDQOELTLGSKYKISVTDIVSILKIISTEKKDSEYTFEV 5580
QY 12 ----- 11
Db 5581 QNDVGRSCKARINVLDLIIPPSFTKKLKMDSIKGSFIDLECIAGSHPISIQWFKDDQ 5640
QY 12 -----THTC-----PP----- 27
| | | | |
Db 5641 EISASEKYKFSFHDNTAFLEISOLEGTDSGTYTCSATNKAGHNQCSGHLTVKREPPYFVEK 5700
QY 28 ----- 27
Db 5701 PQSQDYNPNTRVOLKALVGTA PMTIKWFKNKELHSGAARSVWKDDTSTLSLELFAKAT 5760
QY 28 -----C-----PAP----- 31
| | | | |
Db 5761 DSGTYICQLSNDVGTATSKATLFVKEPPOFIKKPSPVLVRNGQSTTFECQITGTPKIRV 5820
QY 32 ----- 31
Db 5821 SWYLDGNEITATQKHGISFIDGLATFOISGARVENSGTYVCEARNDAGTASCSIELUKVE 5880
QY 32 ----- 31
Db 5881 PPTFIRELKPVEVKYSDVELECEVTGTPPFEVTLKNNREIRSSKKYTLTDRVSFVNLH 5940
QY 32 ----- 31
Db 5941 ITKCDPSDTGEXQCIIVSNEGSGSCSTRVALKEPPSFIKKIENTTTLKSSATFQSTVAG 6000
QY 32 ----- 31
Db 6001 SPPISITWLKDDQILDDEDDNVYISFVDSVATLQIRSDNGHSGRYTCQAKNESGVBCYA 6060
QY 32 ----- 31
Db 6061 FLLVQEPQAIIVEKAKSDVDEKDPMTLECVAGTPELKVWLKDGQIVPSRYFSMFEN 6120
QY 32 ----- 31
Db 6121 NVASFRIQSVMKDQSGYTFKVENDFSSCDAYLRVLDQNI PPPSFTKLTKMDKVLGSS 6180
QY 32 ----- 31
Db 6181 IHWECKVSGSLPISAQWFKDGKEISTSAKYRLVCHERSVSLEVNNLEEDTANYTCCKVSN 6240
QY 32 -----ELLGG 36
| | | | |
Db 6241 VAGDDACSGILTVKEPPSFLVKPGRQQAIPDSTVEFKAILKGTPTPKIKWFKDDVELVSG 6300
QY 37 PSVFL----- 41
| | | | |
Db 6301 PKCFIGLEGSTFNLNYSVDASKTQYTCHTVNDVSDSCTTMLLVTEPPKFKVLEASK 6360
QY 42 ----- 41
Db 6361 IVKAGDSSRLECKIAGSPEIRVWFRNEHELPAOKRYMTFIDSVAVIQMNNLSTEDSGD 6420
QY 42 -----PPP----- 44
| | | | |
Db 6421 FICEAQNAGSTSCSTKVIVKEPPVFSFPPIVETLKNAEVSLECELSGTPPEVVMYKD 6480
QY 45 ----- 44
Db 6481 KRQLRSSKKYKIASKNFHTSIHILNVDTSDIGEYHCKAQNEVGS DTCVCTVKLKEPPREV 6540
QY 45 -----KPKO----- 48
| | | | |
Db 6541 SKLNSLTVVAGEPAELQASIEGAQPIFVQWLKEKEVIRESENIRITTFVENVATLQFAKA 6600

QY 49 -----TLM----- 51
Db 6601 EPANAGYICQIKNDGMEENMATLWLEPAVIVERAGPMVTVGETCTILECKVAGTPEL 6660
QY 52 ----- 51
Db 6661 SVEYKDGKLLTSQKHFSYFNKISSRLRILSVERODAGTYTTFQVQNNVGKSSCTAVDV 6720
QY 52 ----- 51
Db 6721 SDRVPSFTRRLKNTGGVILGASILECKVAGSPISVAMFHEKTKIVSGAKYQTFSDN 6780
QY 52 ----- 51
Db 6781 VCTLQNSLSDMGNYTCVAANVAGSDECEAVLTVQEPFSPVKEPEPLEVLPGKNWTF 6840
QY 52 -----ISRTPEVTCV----- 62
Db 6841 SVIRGTPPFKVNWFRGARELVKGRDCNIYPEDTVAELEFNIDISQGEYTCVVSNNAG 6900
QY 63 ----- 62
Db 6901 ASCTTRLFVKEPAFLKRLSDHSVEPGKSIILESTYTGTLPISVTWKKDGFNITSEKN 6960
QY 63 ----- 62
Db 6961 IVTEKTCILEILNSTKRDAQYSCETIENAGRDVCCALYSTLEPPYFVTEPLEAAG 7020
QY 63 ----- 62
Db 7021 DSVSQQVAGTPEITVSWYKGTKLKRPTEYRTYFTNNVATLVFNKVNINDSGEYTKA 7080
QY 63 ----- 62
Db 7081 ENSIGTASSKTVFRIQERQLPPSFAROLKDIEQTVGLPVLTLCLRLNGSAPIQVCWYRDGV 7140
QY 63 -----VDVSHE----- 68
Db 7141 LLRDDENLOTFSVDNVATLAILQTLDSHGQYSCSASNPLGTASSARLTAREPKKSPFF 7200
QY 69 ----- 68
Db 7201 DIRPVSIDVIAGESADFECHVTGAOPMRITWSKDNKEIRPGNVTITCVGNTPHLRILKV 7260
QY 69 -----DPE 71
Db 7261 KGDSGQYTCQATNDVGKDMCSAQLSVKEPPKFVKKLEASKVAKQGESIQLECKISGPE 7320
QY 72 VK----- 73
Db 7321 IKVSWFRNDSSELHESWKYNSFINSVALLTINEASAFSDSGDYICEAHNGVGDASCSTALT 7380
QY 74 ----- 73
Db 7381 VKAPPVFTQKSPVGVKLGSDVILQCEISGTPPPEVWVKDRQVRNKKFKITSKHFT 7440
QY 74 ----- 73
Db 7441 SLHLINLEASDVGEYHCKATNEVGSDTCSCSVKFEKPPRFVKKLSDTSTLIGDAVELRAI 7500
QY 74 ----- 73
Db 7501 VEGFQPISVVWLKDRGEVIRESENTRISFDINATLQGSPEASNSCKYICQIKNDAGMR 7560
QY 74 ----- 73
Db 7561 ECSAVLTVLEPARIIEKPEMTVTGTGNPFALCECVVTGTPELSAKWFKDGRSADSKHHI 7620
QY 74 ----- 73
Db 7621 TFINVASLKI PCAEMSDKGLYSFEVKNVSGKSNCTVSVHVSRIVPPSFIKRLKOVNAI 7680
QY 74 ----- 73

Db 7681 LGASVVLECRVSGSAPISVGVGFODGNEIVSGPKQSSFSENVCFTNLNLSLLEPSDTGIYTC 7740
QY 74 ----- 73
Db 7741 VAANVAGSDECSAVLTVQEPFSPFQTPDSVEVLTPGMSLTFTSVIRGTPPFKVKWFKGSR 7800
QY 74 ----- 73
Db 7801 LVPGESCNISLEDFVTELELFEVQPLESGDYSCLVTVNDAGSASCTTHLFVKEPATFVKRL 7860
QY 74 ----- 73
Db 7861 ADFSVEGTGSPIVLEATYTGTPPISVSWIKDEYLIQSQCSCSITWTEKSTILEILESTIED 7920
QY 74 ----- 73
Db 7921 YAOVSCLIEAGODICEALVSVLEPPYFIEPLEHVEAVIGEPAATLQCKVDGTPEIRISW 7980
QY 74 ----- 73
Db 7981 YKEHTKURSAPAYKMQPKNNVASLVINKVDHSDVGEYSCRADNSVGAVASSAVLVIKERK 8040
QY 74 -----FNTVYDGV----- 81
Db 8041 LPPFFARKLKDVBHETLGFPPVAFECRINSEPLQVSWYKDGVLKKDDANLQTSFVHNVATL 8100
QY 82 ----- 81
Db 8101 QILQTDQSHIGQXNCASNPLGTASSAKLILSEHVEPPFFDLKPVSDVLDALGESGTFKC 8160
QY 82 ----- 81
Db 8161 HVTGTAPKITWAKDNREIRPGGNKMTLVENTATLVLVKVGKDAGQYTCYASNIAGKD 8220
QY 82 ----- 81
Db 8221 SCSAHLGVQEPFRIKKLEPSRIVKQDEFTRYECKIGSGPEIKVLWKDETEIQESSKFR 8280
QY 82 -----EVHNA----- 86
Db 8281 MSFVDSVAVLEMHNLNLSVEDSGDYTCEAHNAAGSASSSTSLKVKPEPIFRKKPHIETLKG 8340
QY 87 ----- 86
Db 8341 ADVHLECELOGTPPFPHVSWYTKDKRELKSGKKYKIMSENFLTSIHLNVDAADIGEYOCKA 8400
QY 87 -----KTKPR----- 91
Db 8401 TNDVGSOTCVGSIALKAPPRFVKKLSDISTVVGKEVQLQTTIEGAEPISVWVKDKGEIV 8460
QY 92 ----- 91
Db 8461 RESDNIWISYENIATLQFSRVEPANAGKYTCQIKNDAGMQECPATLSVLEPATIVKEPE 8520
QY 92 ----- 91
Db 8521 SIKVTTGDTCTLECTVAGTPELSTKWKFGDKELTSDNKKYKISFKNVSGLKIINVAFSDS 8580
QY 92 ----- 91
Db 8581 GYVSFEVQNPVGKDCTASLQVSDRTVPPSFTRKRLKETNGLSGSVVWECKVYGPPIV 8640
QY 92 -----EE----- 93
Db 8641 SWFHEGNEISSGRKYQTLTDNTCALVTNMLEESDSDGYTCIATNMAGSDECSAPLTVRE 8700
QY 94 ----- 93
Db 8701 PPSEVQKPDMDVLTGTNTVFTTSIVKGTPPFSWFKGSSSELVPGDCRNVSLSDVAELE 8760
QY 94 ----- 93

Db 8761 LFDVDTQSSEYTCIVSNEAGKASCTTHLYIKAPAKFVRLNDYSIEKGPLILEGTFGTG 8820
QY 94 -----QYN----- 96
Db 8821 TPIPSVTWKNGINVTSPQRNITITTEKSAILEIPSSSTVEDAGQYNCYIENASGRDSCSA 8880
QY 97 -----STYRV----- 101
Db 8881 QILILEPPYFVKOLEPVKVSVDGSASLOCLAGTPEIGVSWYKGDTKLRPTTTYKMHFRN 8940
QY 102 ----- 101
Db 8941 NVATLVFNQVDINDSCEYICKAENSVEVSASTFLTQVQOKLPPSPSROLRDVQETVGLP 9000
QY 102 -----VS----- 104
Db 9001 VVFDCAISGEPIISWSYKDGKPLKDSNNVQTSFLDNTATLNIKFTDRSLAGQYSCSTATN 9060
QY 105 ----- 104
Db 9061 PIGSASSARLILTEGKNPPFFDIRLAPDVAVVGESADFECHVTGTQPIKVSNAKDSREI 9120
QY 105 -----LTVLHOD-----WLNKGE----- 117
Db 9121 RSGGKQIISYLENSAHLTVLKVKGDSGQYTCYAVNEVGKDSCTAQLNIKERLIPPSFTK 9180
QY 118 ----- 117
Db 9181 RLSETVEETEGNSFKLEGRVAGSQPITVAVYKNNIEIQTSNCEITFFKNNTLVQVRKAG 9240
QY 118 -----YKCKVSN----- 124
Db 9241 MNDAGLYTCKVSDNAGSALCTSSIVIKPKPPVDFDHLTPVTVSEGEYVQLSCHVQSGE 9300
QY 125 ----- 124
Db 9301 PIRIOWLKAGREIKPSDRCSFSFASCTAVLELRDVAKADSGDYVCKASNVAGSDTTKSKV 9360
QY 125 ----- 124
Db 9361 TIKDRPAVAPATKAAVDGRLEFFVSEPOSIRVVEKTATFIKVGCDPIPNNVKTGKWR 9420
QY 125 ----- 124
Db 9421 QLNQGRVFIHQGDGALEIRDTTKDGLYRCVAFNEHGEIESNVNLQVDERKKQEKI 9480
QY 125 ----- 124
Db 9481 EGDRLMLKKTPIILKKGAGEEEDIMELLKNVDPKVEKYARMYGITDFRGLLQAFELL 9540
QY 125 ----- 124
Db 9541 KQSQEETHRLIEIERSERDEKEFEELVSFIQORLSQTEPVTLIKDIENQTVLKNDNA 9600
QY 125 ----- 124
Db 9601 VFEIDIKINPEIKLSWYKTEKLEPSDKFETSIDGRHRTLKVKNCQLKDQGNRYLVCGP 9660
QY 125 ----- 124
Db 9661 HIASAKLTVIEPAWBRHLQDVTLKEGQCTMTQCFSPVNVKSEWPRNGRIILKPOGRHKTE 9720
QY 125 ----- 124
Db 9721 VEKVKHLTIADVRAEDQGYTKVEDLETSAELEIAEPIQFTKRIONIVVSEHQSATF 9780
QY 125 ----- 124
Db 9781 ECEVSFDDAIVTWYKGTLETESQYKFNDRGCRHYMTIHNVTPDDEGVYSVIARLEPRG 9840
QY 125 ----- 124
Db 9841 EARSTAELYLTTKEIKLELKPDPIDPSRVPIPTMPIRAVPPEEIPPVVAPPILPLLTPE 9900

QY 125 ----- 124
Db 9901 EKKPPPRIEVTKKAVKDKAKKVAKPKEMTPREEIVKKPPPTTLIPAKAPEIIDVSSK 9960
QY 125 ----- 124
Db 9961 AEEVKIMTITRKKEVQKEAVEYKQAVHKRVFIESFEEPYDELEVETPEPQY 10020
QY 125 ----- 124
Db 10021 YEEPDDEYEEKVAKEVHEEWEEDFEBGQYERYEREBGYDEGEWEWEAYOEREVIQVO 10080
QY 125 ----- 124
Db 10081 KEVYEESHERKVPKVPKKAAPPKPKVKKPVIEKIEKTSRRMEBEKVQVTKVPEVSKKI 10140
QY 125 ----- 124
Db 10141 VPQKPSRTPVQEEVIEVKVPAVHTKKMVISEKMFASHTEESESVTVPEVQKEIVTEBK 10200
QY 125 ----- 124
Db 10201 IHVAVSKRVPPPKVPPELPEKPAPEEAVPPIPKKVEPPAPKVPPEKPKVPPEKKPVPV 10260
QY 125 ----- 124
Db 10261 PKKEPAAPPKVPPEKPKVPPEEKIPVPVAKKKEAPPKVPPEVQKRVVTEEKITIVTORRE 10320
QY 125 ----- 124
Db 10321 SPPPAVPEIPKPKVPPEERKVPPEEVEVPPPKVPALPKVPPEEKVAVVPVAKKAPPP 10380
QY 125 ----- 124
Db 10381 RAEVSKKTVEEKRFVABEELSFVAPQVREVTRHEVSABEEESYSVEEGVSI SVYREEE 10440
QY 125 ----- 124
Db 10441 REEEBAEVEYEVMEPEEYVVEEKLHIISKRVEAEPAEYTEROEKKIVLKPAPAKIE 10500
QY 125 -----KALPAPI----- 131
Db 10501 EPPPAKVPAPKKIVPEKKVPAPVPKKEKVPPEKPKVPPEEPPKVPPEKVPKVPKKEEPLP 10560
QY 132 ----- 131
Db 10561 AKVTERHMOITQEEKVLVAVTKKEAPKARVPEEPKRAVPEEKVLKPKKREEPPAKVT 10620
QY 132 ----- 131
Db 10621 EFRKRVKKEKVSIEAPKREPOPIKEVTIMEKERAYTLEEEAVSVQREEEYEEYDY 10680
QY 132 ----- 131
Db 10681 KEFEYEPTDYDQYEEYEREYERYESHEEYITEPEKPIPVKVPPEEPVPTKAPPAK 10740
QY 132 ----- 131
Db 10741 VLKKAPEEKVPVPIPKKLKPPPKVPPEEKVFEKIRISITTKREKQVTEPAKAPVPMK 10800
QY 132 ----- 131
Db 10801 PKRVVAEEKVPVRKEVAPPVVRVPEVPKLEPEEVEVAFEEVVVTHVEEYLVBEEEYIHHE 10860
QY 132 ----- 131
Db 10861 BEFITEEVWPVIVKVPPEVPRKVPPEEKKVPVPKKEAPPKVPPEVPPKPKPEEKVPVLI 10920
QY 132 ----- 131
Db 10921 PKKEKPPPAKVPPEVPPKVPPEEKVPVPVPPKVEAPPKVPPEKPKVPPEKVPVAPKKV 10980

Qy 132 ----- 131
Db 10981 EAPPAKVPVPPKLIPEKKKTPVPKVEAPPKVPKREPVPVVALPQEEVLFEEI 11040
Qy 132 ----- 131
Db 11041 VPEEVLPEEEVLPEEEVLPEEEIPEEEVPEEEVPEEEVPEEEV 11100
Qy 132 ----- 131
Db 11101 LPEVKVPVPVPPVPEIKKVKTEKKVVPKKEAPPKVPKVPKVEEKRIILPKEEVL 11160
Qy 132 ----- 131
Db 11161 PVEVTEPEEPISEEEIPEEPPSTEVEEVPVPRVPEVIKAVPEATPVPKVEAPPA 11220
Qy 132 ----- 131
Db 11221 KVSKKIPEKVPVPVORKEAPPKVPKVPKVEKKVLVPKKEAVPPAKGRTVLEEKVSV 11280
Qy 132 ----- 131
Db 11281 AFROEVVVKERLEVEVAEVEEIPPEEPEHEVEYFEEGFEHEVEEPIKLEQHRVEEH 11340
Qy 132 ----- 131
Db 11341 RVEKVRVIEVPEAEVEVEFEKPKAPKGPSEIKLIIPPKKPPTKVVPKPEPPKAPVPEV 11400
Qy 132 ----- 131
Db 11401 KKIIVEEKVRVPEBPRVPPTKVDPVLPKVEVPEKKVPVPPAKKPEAPPKVPKPEV 11460
Qy 132 ----- 131
Db 11461 PEKKVPPPPKKPEVPTTKVPEVKAAPVEKKVPEAIPPKPEPPPEVPEAPKEVVPKK 11520
Qy 132 ----- 131
Db 11521 VPAAPPKKPEVTPVKVPEAPKEVVPKVPVPPPKKPEVPTKVPEVPAVPEKKVPEA 11580
Qy 132 ----- 131
Db 11581 IPPKPEPPPEVEEVALEEPPAEVVEPEAPAPQVTPVPKVPKPEKAPVAKK 11640
Qy 132 ----- 131
Db 11641 PELPPVKVPKEVVPKVPVLPVVPKPEAPPKVPKEVVPKVPKVPKVPKVPKVP 11700
Qy 132 ----- 131
Db 11701 KVPEVPPKVPLEKPAVPVPRAESPPPEVVEEPEEIAPEEIEAPEEKPVPVAAEEPE 11760
Qy 132 ----- 137
Db 11761 VPPPAVEPKKIIPEKKVPVVIKPEAPPKPEPEPEKVIERPCLKRPPPPPPAPPKEDV 11820
Qy 138 ----- 137
Db 11821 KEKIFOLKAIPKKVPKVPQVPEKVELTPLKVPGEKKVKRLLPERKPEPEEVLKSVL 11880
Qy 138 ----- 137
Db 11881 RKRPEEPEKVPKLEKVPKPAVPEPPPPKPEVEVEVPTVKRERKIPETKVPKIPA 11940
Qy 138 ----- 137
Db 11941 IPLPAPEPKPEAEVTKIKPPVPEPTIAAPVTPVVGKAEAKAPKEAAKPGPI 12000
Qy 138 ----- 137
Db 12001 KGVPKTPSPAERRLRPGSGGKPPDEAPFTYQLKAVPLKFVKEIKDIILTESEFVG 12060
Qy 138 ----- 137

Db 12061 SSAIFECVLSFSTALTTHMKDGSNIRESPKHRFIADGKDRKLHIIDVQLSDAGEYTCVLR 12120
Qy 138 ----- AKQP ----- 142
Db 12121 LGNKEKTSTAKLVVEELPVRFVKTLEEBVTVMKGPLYLSCELNKRVDVVRKDKGI VVE 12180
Qy 143 ----- 142
Db 12181 KPGRIVPGVIGIMRALTINDADDTAGTYTIVTENANNLECCSVKVVVEVIRDWLVKPIR 12240
Qy 143 ----- 142
Db 12241 DOHVKPKGTALFACDIANDTPNIKWFKGYDEIPAEPNDKTEILRDGNHLYLKIKNAMPED 12300
Qy 143 ----- 142
Db 12301 IAEYAVEIEGKRYPAKLTGGERVELLKPIEDVTIYERESAFDAEISEADIPGQWLKG 12360
Qy 143 ----- 142
Db 12361 ELLRPSPTCEIKAEKGKRFLTLRKVKLDQAGEVLYQALNATTAITLTKEIELDFAVPLK 12420
Qy 143 ----- 142
Db 12421 DVTVPERQARFECVLTREANVINSGPDI IKSSDKFDI IADGKKHILVINDSQFDEGV 12480
Qy 143 ----- 142
Db 12481 YTAIEVEGKTSARLFTVGTIRLKFMSPLEDOTVKEGETATFVCELSHERKMHVWFKNDAKL 12540
Qy 143 ----- 142
Db 12541 HSTRVTLISSEGTHKLEMKVETLDDISQIRAOVKELSSAQKLVLEADPTFTVKLHDKT 12600
Qy 143 ----- 142
Db 12601 AVEKDEITLKCEVSKDVPVKWFKDGEIIVSPKYSIKADGLRRILKIKKADLKDKGEVVC 12660
Qy 143 ----- 142
Db 12661 DCGTDTKANVTVEARLIKVEKPLYGVFVGETAHTAFIEISEPDVHGQWLKGQPLTAS 12720
Qy 143 ----- 142
Db 12721 PDCEIIEDGKKHILILHNCQLGTMGEVSFOAANAKSAANLKVKEPLPLIETPLSDVKVFE 12780
Qy 143 ----- REPO ----- 146
Db 12781 KDEAKFECEVSREPKTFRWLKGTQETITGDRFELIKDGTKHSWIKSAAFEDEAKYMFEA 12840
Qy 147 ----- 146
Db 12841 EDKHTSGKLIIEGIRLKLFTPLKDVTAKEKESAVFTVLSHDNIRVKWFKNDQRLHTRS 12900
Qy 147 ----- 146
Db 12901 VSMQDEGKTHSITFKDLSIDDTSQIRVEAMGMSSEAKLTVLEGDPTFTGKLQDTGVKED 12960
Qy 147 ----- 146
Db 12961 EVILQCEISKADAPVKWFKDGEIKPKNAVIKADGKKRMLILKALKSDIGQYTCDCGT 13020
Qy 147 ----- 146
Db 13021 DKTSGKLDIEDREIKVRPLHSHVEVMETETARFETIEISDDOIHANWKLKGEALLQTPDCE 13080
Qy 147 ----- 146
Db 13081 IKEEGKIHSVLHNCRLDQDQGVDFQAANVKSAAHLRVKPRVIGLLRPLKDVTVTAGETA 13140
Qy 147 ----- 146

Db 13141 TFCELSYEDIPVEMYLGKKLEPSPDKVVRSEKGVHTLTLRDVKLEDADEVOLTAKDFX 13200
QY 147 ----- 146
Db 13201 THANLFVKEPPVETKPLEDOTVEEGATAVLECEVSRENAKVWFKNGTILASKKYEIV 13260
QY 147 ----- 146
Db 13261 ADGRVRKLVTHDCTPEDIKYTCDAKDKTSCNINVPPHVEFLRPLTDLQVREKEMARF 13320
QY 147 ----- 146
Db 13321 ECELSRENAKVKWFKDGAIEIKGKYDIISKGAVRILVINKCLLDEAEYSCEVRTARTS 13380
QY 147 ----- 146
Db 13381 GMLTVLEEEAVFTKNLANIEVSETDIKLCEVSKPAGAEVIWYKGBDEIETGRYEILTE 13440
QY 147 ----- 146
Db 13441 GRKRILVIONAHLEDAGNYCNRLPSSRTDGKVKVHELAAEFISKPNLEILEGEKAFFVC 13500
QY 147 ----- 146
Db 13501 SISKESFPVQMKRDKTLESGDKYDIADGKKRVLVVKDATLQDMGTYYVMVGAARAAH 13560
QY 147 -----V 147
Db 13561 LTWIEKLRIVVPLKDRVRKEQEVFENCEVNTGAKAKWFRNEEAIFDSSKYIILQKOLV 13620
QY 148 YTL----- 150
|||
Db 13621 YTLIRDAHLDDQANYNSLTNRGENVKSAANLIVEEEDLRIVEPLKDIEIEMKKSIVF 13680
QY 151 ----- 150
Db 13681 WCKVNRNLNVLKWTNGEEVFPDNRVSYRVDYKXHLITIKDCGFPDEGEYIVTAGODKSV 13740
QY 151 ----- 150
Db 13741 AELLIIIEAPTEFVEHLEDQTVTEFDDAVFSCQLSREKANVWYRNGREIKGKKYKFEKD 13800
QY 151 ----- 150
Db 13801 GSIHRLIIKDCRLDDEYACGVEDRSRRLFVEEIPVEIIRPPQDILEAPGADVFLA 13860
QY 151 ----- 150
Db 13861 ELANKDVEQWLRNNMVVQGDQKQMMSEKTHRLQICDIKPRDQGEYRFAKDKEARAK 13920
QY 151 ----- 150
Db 13921 LELAAAPKIKTADQDLVDVVGKPLTWVPVDPAYPKAAEWFKENEPSTKTIDTTAEQTS 13980
QY 151 ----- 150
Db 13981 FRILEAKGDKGRYKIVLQNHKGAEFNLKVIDVPGVPRNLEVTETEDGEVSLAWEPP 14040
QY 151 ----- 150
Db 14041 LTDGSKIIGYVVERRDIRKKTWVLATDRAESCEFTVTGLQGGVEYLFVRSARNRVGTG 14100
QY 151 ----- 150
Db 14101 EPVETDNPVEARSKYDVGPPPLNVITIDYNREGVSLTWEPPEYDGAETINVYIELRDKT 14160
QY 151 ----- 150
Db 14161 SIRWDTAMTVRAEDLSATVTDVVEGOEYSFRVRAQNRIGVGKPSAATPEVKVADPIERDS 14220
QY 151 ----- 150
Db 14221 PPNVLTSSDOTQSSVOLKKEPPLKDGSGPILGYIIRCEEGKDNWJRCNMKLVPELTYKV 14280

QY 151 ----- 150
Db 14281 TGLEGNKYLRYSAENKAGVSDPSEILGLPTADDAFVEPTMDLSAFKDCGLEIVPNPIT 14340
QY 151 ----- 150
Db 14341 ILVPSTGYPRPTATWCFGDKVLETGDRVKMKTLTSAVAELVISPERSDKGIYTLKLENRV 14400
QY 151 ----- 150
Db 14401 KTISGEIDVNVIAIRPSAPKELKFGDITKDSVHLTWEPDDDGSPLTGYVVEKREYSRKT 14460
QY 151 ----- 150
Db 14461 WTKVMDFTDLETTVPDLVQCKEYLFKVCARNKCGGEPAYVDEPVNMSTPATVPDPEN 14520
QY 151 ----- 150
Db 14521 VKWRDRTANSIFLTWDPKNDGSGRIKGYIVERCPRGSDKWACGEPVAETKMEVTGLEE 14580
QY 151 ----- 150
Db 14581 GRWYAYRVKALNRQGASKPSRPTTEIQAVDQEAPEIFLDVKLLAGLTVKAGTKIELPAT 14640
QY 151 ----- 150
Db 14641 VTGKPEPKITWTWKADMILKQDKRITIENTIENVPKSTVTIVDSKRSDTGTYYIIIEAVNVCGRAT 14700
QY 151 ----- 150
Db 14701 AVVEVNVLDKPGPAAFDITDVTNESCLLTNPPRDDGSKITNYVVERRATDSEVWHKL 14760
QY 151 ----- 150
Db 14761 SSTVKDTNFKATKLIPNKEYIFRVAENMYGVGEVQASPIATAKYQDPPGPTRLEPSD 14820
QY 151 ----- 150
Db 14821 ITKDAVTLTWCEPDDDDGSGPITGYWVERLDPTDKWRCNKMPVKDITTVRVKGLTNKKY 14880
QY 151 ----- 150
Db 14881 RFRVLAENLAGPKSKSTEPILIKDPIPPWPPGKPTKDVGKTSVRLNWTKEHDGGA 14940
QY 151 ----- 150
Db 14941 KTESYVIEMLKTGTDENVRVAECVPTTQHLLPGLMEGQEYSFRVRAVNKAGESEPEPSD 15000
QY 151 ----- 150
Db 15001 PVLCREKLYPPSPRRWLEVINITKNTADLKWTVPKDGSGSPITNYIVEKRDVRKRCQWTV 15060
QY 151 ----- 150
Db 15061 DTTVKDKCTVPTLEBSLYFRVAENAIGOSDYTEIEDSVLAKDFTTTPGPPYALAVV 15120
QY 151 ----- 150
Db 15121 DVTKRHRVDLKWBPKNKNDGGRPIQRYVIEKKERLGTNRWVAGKTAGPDCNFRVTDVIEGTE 15180
QY 151 ----- 150
Db 15181 VOQVRAENEAGVGHPSPEITELSIDPTSPSPPLDLHVTADGRKHIAIAWKPPKNGG 15240
QY 151 ----- 150
Db 15241 SPIIGYHVEMCPVGTGKWRVNSRPIKDLKFRVEGVVDPKEYVLRVRAVNAIGVSEPS 15300
QY 151 ----- 150
Db 15301 ISENVAKDPDCKPTIDLETHDIIVIEGKLSIPVFPRAVPVTVSNHKDGKEVKASDRL 15360

QY 151 ----- 150
Db 15361 TMKNDHISAHEVPKSVRADAGIYITITLENKLSATASINVKIGLPGCKDIKASDITK 15420
QY 151 ----- 150
Db 15421 SSCKLTWEPPEFDGGTPIHLVLERREAGRRTYIPVMGENKLSWTVKOLIPNGEYFRV 15480
QY 151 ----- 150
Db 15481 KAVNKVGGGEYIELKNPVIADQPKQPPDPVVEVHNPTAEAMTITWKPLYDGGSKIMG 15540
QY 151 ----- 150
Db 15541 YIIETIAKGEERNKRCNEHLVPILTYTAKGLEBGEYQFRVRAENAAIGSEPSRAPPTK 15600
QY 151 ----- 150
Db 15601 AVDPIDAPKVILRTSLEVKRGDEIADASISGSPYPTITWIKDENVIVPEEIKKRAAPLV 15660
QY 151 ----- 150
Db 15661 RRRKGEVQEEPPFVLPLTQRLSTDSNKKGESQLRVRDSLDPDHGLYMIKVENDHGIAKAP 15720
QY 151 ----- 150
Db 15721 CTVSVLDTGPPINFVFEDIRKTSVLCKWEPPLDDGSGSEIINYTLEKKDKTKPDSEWIVV 15780
QY 151 ----- 150
Db 15781 TSTLRHKYSVTKLIEGKEYLFRVRAENRFGPGPCVPLVAKDPFGPPDAPDKPVED 15840
QY 151 ----- 150
Db 15841 VTSNMLVKWNEPKDNGSPILGIWLEKREVNSTHWSRVNKSLLNALKANVDGLLEGLTYV 15900
QY 151 ----- 150
Db 15901 FRVCAENAGPGKFPSPDKTAHDPISPGPPPIPRVTDTSSTIELEWEPFANGGEI 15960
QY 151 ----- 150
Db 15961 VGYFVDKQLVGTNEWSRCTEKMIKROYTVKEIREGADYKLRVSANAAEGGPGETQPV 16020
QY 151 ----- 150
Db 16021 TVAEPPQPPAVELDVSVKGGIOMAGKTLRIPAVTVGRPVPTKVWTKKEGELDKDRVID 16080
QY 151 ----- 150
Db 16081 NVGTSKSELIKDALKRDHGRVITATNSCGSKFAAARVEYFDPGPVLDLKPVVTRKMC 16140
QY 151 ----- 150
Db 16141 LLNWSDEDDGGSEITGTIIERDKAKMHTWROPIETERSKCDITGLLEGOYKFRVIK 16200
QY 151 ----- 150
Db 16201 KFGGPPVEIGPILAVDPLGPPTSPERLTYTERTKSTITLDWKEPRNGSGSPIQGYIEK 16260
QY 151 ----- 150
Db 16261 RRHDKPDFERNKRLCTTSFLVNLDELHOMYEFVRKAVNEIGESEPSLPINVVIOQDEV 16320
QY 151 ----- 150
Db 16321 PPTIKRLSVRGDTIKVKAQEPVHIPADVTGLPMPKLEWSKNETVIEKPTDALQITKEEV 16380
QY 151 ----- 150
Db 16381 SRSEAKTELSIPKAVREDKGTYYVTASNRLGVSFRNVHVEYDRPSPRNLAVIDIKAES 16440
QY 151 ----- 150

Db 16441 CYLTWADPLDNGSGSEITHYVIDKRDASRKKAWEVINTAVAKRYGIWKLPNGQYFRV 16500
QY 151 ----- 150
Db 16501 RAVNKYGISDECKSKVVIQDPYRLPGPGPKVLARTKGSMLVSWTPPLDNGGSPITGY 16560
QY 151 ----- 150
Db 16561 WLEKREGSPYWSRVSRAPITKVGLKGVEFNVPRLLEGVKYQFRAMAINAAGIGPPSEPS 16620
QY 151 ----- 155
|:::|
|:::|
Db 16621 DPEVAGDPIPPGPPSPCEVKDKTKSSISLGWKPPAKDGGSPIKGYIVEMQEETTDWKR 16680
QY 156 ----- 155
Db 16681 VNEPDKLITTCVCPVNLKELRYRFRKAVNEAGESESDTTGEIPATDIOEEPEVID 16740
QY 156 ----- 155
Db 16741 IGAQDCLVCKAGSQIRIPAVIKGRPTPKSSWEFDGKAKKAMKDVHDIPEDAQLETAENS 16800
QY 156 ----- 155
Db 16801 SVIIIECKRSHTGKYSITAKNAGQKTANCRVKVMDVPGPDKLVSDITRGSCRLSWK 16860
QY 156 ----- 155
Db 16861 MPDDGGDRIGYVIEKRTIDGKAWTKVNPDCGSTTFVVPDLJSEBQQYFFRVAENRFGI 16920
QY 156 ----- 155
Db 16921 GPPVETIQTARTDPIYPPDPPIKLKIGLITKNTVHLSWPKPKNDGSGSPVTHYIVECLAW 16980
QY 156 ----- 155
Db 16981 DPTGTKEANROCNKRDRVEELQFTVEDLVEGGEYFRVKAVNAAGVSKSPSATVGPCDQR 17040
QY 156 ----- 155
Db 17041 PDMPPSIDLKEFMEVEBGTNVNIVAKIKGVFPPTLTWFKAPPKPNKEPVLVDYHVNKL 17100
QY 156 ----- 155
Db 17101 VDDTCTLVIPQSRSDTGLYTTITAVNNLGTASKEMRLNVLGRPGPPVGIKPEVSADQ 17160
QY 156 ----- 155
Db 17161 MTLSPFPKDDGSKITNYVIEKREANRKTWVHVSSEPKECTYTIKLEGEHYVFRIMA 17220
QY 156 ----- 155
Db 17221 QNKYIGEPLDSEPTARNLFSVPGADKPTVSSVTRNSMTVNWEEPEYDGGSEVTVYWL 17280
QY 156 ----- 155
Db 17281 EMKDTTSKRWKVRNRDPIKAMTLGVSYKVTGLIEGSDYQFRVAINAAGVGPASLPSDPA 17340
QY 156 ----- 155
Db 17341 TARDPIAPGPPPKVTDWTWKSSADLEWSPPLKDDGSKVTGYIYVEYKEEGKEWEKCKDK 17400
QY 156 ----- 155
Db 17401 EVRGTKLVVTKLEGAFFYKFRVSAVNIAGIGEPEGVEDVIEMKDRLVSPDLQDASVRDR 17460
QY 156 ----- 155
Db 17461 IVVHAGVIRIIAYVSGKPPPTVTWNNMERTLPQEATITETAISSSVIKNQCRSHOGVY 17520
QY 156 ----- 155

Db 17521 SLLAKNEAGERKKTIIVDVLDPGPGVGTPLAHLNLTNESCKLTWFSPEDDGGSPITNYVI 17580
QY 156 ----- 155
Db 17581 EKRESORRAWTPYTYVTQRONATVOGLIOQKAYFFRIAAENSGMGPFVETSEALVIREP 17640
QY 156 -----ELTKNOVSLT----- 165
Db 17641 ITVPERDELEVKEVTKNTVLTWNPKYDGGSEIINYLESRLIGTEKFKVTNDNLLS 17700
QY 166 -----CLV- 168
Db 17701 RYTVKGLKRGDYERYRSANIVGOGKPSFCTKPTCKDELAPPTLHLDLFRDKLIRVG 17760
QY 166 -----CLV- 168
Db 17761 EAFALTGRYSGKPKPKVSKFKEADVLEDDRTHIKTTPATLALEKTKAKRSOGSKYCVV 17820
QY 169 -----KGF----- 171
Db 17821 ENSTGSRKGFQCVNVVDRPGPPVPGVSFDEVTKDYMVISKKPLDDGGSKIITNYIEKKE 17880
QY 172 ----- 171
Db 17881 VGKDVMPVTSASAKTTCCKVSKLLEGKDYIFRIHAENLYGIDPLVSDSMKAKDRFRVPD 17940
QY 172 ----- 171
Db 17941 APDQPIVTEVKDSALVTWNKPHDGGKPTITNYILEKRETMKRWARVTKDPIHPTKFRV 18000
QY 172 ----- 171
Db 18001 POLLEGQYEFVRSARENEIGDPPSPKPVFAKDPFAKPPVPVPEAIDTTCNSVDLTW 18060
QY 172 ----- 171
Db 18061 QPPRHGGSKILGYIYEVQKVGDEWRANHTPESCPETKYKVTGLRDGQTYKFRVLAVN 18120
QY 172 ----- 171
Db 18121 AGESDPAHVPEPVLVKDRLEPPELILDANMAREQHIKVGDTLRLSAILIKGVFPFKVTWK 18180
QY 172 ----- 171
Db 18181 KEDRAPTKARIDVTPVGSKEIRNAAHEDGGIYSLTVENPAGSKTVSVKVLVLDKPGPP 18240
QY 172 ----- 171
Db 18241 ROLVSEIRKDCSYLTWKEPLDDGGSVITNYVVERRDVAASQWSPLSATSKKSHFAKHL 18300
QY 172 ----- 171
Db 18301 NEGNQYLFRVAENQYGRGPFVETPKPKIKALDPLHPPGPKDLHHVDVDKTEVSLVWNKP 18360
QY 172 ----- 171
Db 18361 DRDGGSPITGYLVEYQEGTDWIKFKVTNLECVVTGLQOGKTYRFRVKAENIVGLGLP 18420
QY 172 ----- 171
Db 18421 DTTPIEQEKLVPSPVELDVKLIBGLVVVAGTTVRFPAILIRGVVPVPTAKWTTDGESEIKT 18480
QY 172 ----- 171
Db 18481 DEHYTVETDNFSSVLTIKNCLRDTEGYQITVSNAAGSTVAVHLTVLDVPGPTGPINI 18540
QY 172 ----- 171
Db 18541 LDVTPHEMTISWQPKDKDGGSPVINYIVEKQDRKDTWGVSSGSKTKLKIPHLQKGE 18600
QY 172 ----- 171
Db 18601 YVFRVRAENKIGVGPPLDSTPTVAKHKFSPSPGKPVVTDITENAAVTSWTLPSKDGGS 18660

QY 172 ----- 171
Db 18661 PITGYMERREVTGKVRVNKTPIADLKPRVTGLYEGNTYEFVRVFAENLAGLSKPSSTD 18720
QY 172 ----- 171
Db 18721 PIKACRPIKPPGPPINPKLKDKSRETADLVWTKPLSDGGSPILGYVVECOKPGCTAQNRI 18780
QY 172 ----- 171
Db 18781 NKDELIROCAFVRVPLIEGNEYRFRIRIKAANIVGEGBRELAESVIAKDILHHPPELVDVT 18840
QY 172 ----- 171
Db 18841 CRDVTVRVGQITIRILARVKGRPEPDITWTKEGKVLVREKRVLDLIQDLPRVELQIKEAVR 18900
QY 172 ----- 171
Db 18901 ADHGKVIISAKNSSGHAQSAIVNVLDPRGPCQNLKAVTNVTKENCTISWENPLDNGSEI 18960
QY 172 ----- 171
Db 18961 TNFIVEYRKPNOKGWSIVASDVTKRLIKANLLANNEYFRVCAENKVGVTIETKPTIL 19020
QY 172 ----- 171
Db 19021 AINPIDRPGEPENLHIADKGTFTVYLKRRPDYDGGSPNLSTHVERRLLKGGDDWVRHKG 19080
QY 172 ----- 171
Db 19081 SIKETHYMDRCVENQIYEFVRQTKNEGSDWVKTEVVVKEDLQKPVLDLKSGLTV 19140
QY 172 ----- 171
Db 19141 KAGDTIRLEAGVRGKPFPEVAWTKDKDATDLTFRSPRVKIDTRADSKFSLTAKRSDGK 19200
QY 172 ----- 171
Db 19201 YVVATATAGSFVAVATVNVLDKPGVNRNLKIVDVSSDRCTVCWDPPEDDGGCEIQNYIL 19260
QY 172 ----- 171
Db 19261 ECKETKRMVWSTYSATVLTPTGTVTTRLIEGNEYIFRVRAENKIGTGPPTESKPIAKTKY 19320
QY 172 ----- 171
Db 19321 DRPGRPDPPEVTKVSKEEMTVVWNPPEYDGGKSITGYFLEKKEKHSTRVWVPVKNKSAIPER 19380
QY 172 ----- 171
Db 19381 RMKVQNLPLPDHEYQFRVKAENEIGIGEPSLPSRPVVVAKDPIEPPGPTNFRVVDTTKHSI 19440
QY 172 ----- 171
Db 19441 TLGWKGVYDGGAPIIGYVVENMRPKIADASPDGKWRKNAAAQLVRKEFTVTSLDENOEY 19500
QY 172 ----- 171
Db 19501 EFRVCAQNOVGIRPAELKEAIKPKKEILEPPEIDLDASMRKLVIVRAGCPIRLFAIVRGR 19560
QY 172 ----- 171
Db 19561 PAKVYTRKRGIDNVVRKGOVDLVTMAFLVTPNSTRODSSGKYSLLVNPAGEKAVFVNV 19620
QY 172 ----- 171
Db 19621 RVLDTPGPVSDLKVSADVTKTSCHVSWAPPENDGGSQVTHYIVEKREADRKTWSTVTPEVK 19680
QY 172 ----- 171
Db 19681 KTSFHVWNLVPGNEYFRVTVAVNEYGPGVPTDVPKPVLASDPLSEDPDRPKLEVTEMTKN 19740

Qy 172 ----- 171
Db 1741 SATLAWPLRLDGGAKIDGYITSYREEEQPADRWTEYSVVKDLSLVVTGLKEGKKYFRV 19800
Qy 172 ----- 171
Db 19801 AARNAGVSLPREAGVYEAKEQLLPKILMPEOITIKAGKKLRIEAHVYKGPHTCKWK 19860
Qy 172 ----- 171
Db 19861 KGEDEVTSHLAVHKADSSILLIKDVTBKDSGYISLTAENSSGTDQIKVVMMDAPG 19920
Qy 172 ----- 171
Db 19921 PPQPPFIDISDADACSLSWHIPLEDGGSNITNYIVEKCDVSRGDWVTALASVTKTSRV 19980
Qy 172 ----- 171
Db 19981 GKLPGQEYIFRVAENRFGISEPLTSPKMWAOFPFGVPSEPKNARVTKNKDCIFVWD 20040
Qy 172 ----- 171
Db 20041 RPDSDGSGPIIGYLLIERKERNLLWVKANDTLVRSTEYPCAGLVEGLEYSFRIYALNKAG 20100
Qy 172 ----- 171
Db 20101 SSPPSKPTEYVARMVPDPPCKPEVIDVTKSTVSLIWARPKHDGSGKIIGYFVEACKLPG 20160
Qy 172 ----- 171
Db 20161 DKWVRTAPHOIQEYETATGLEEKAQYQFRAIARTAVNISPPSESDPVTILAENVPP 20220
Qy 172 ----- 171
Db 20221 RIDLSVAMKSLITVKAAGTNVCLDVTFGKPMPTVSWKKDGTLLKPAEGIKMAMORNLC TL 20280
Qy 172 ----- 171
Db 20281 ELFSVNRKDSGYTITAENSSGKSATIKLVLDKPGPASVKINKMYSDRAMLSWEPPL 20340
Qy 172 ----- 171
Db 20341 EDGSGEITNYIVDKRETSRPNWAQVSATVPITSCSVEKLIIEGHEYQFICAENKYGVDP 20400
Qy 172 ----- 171
Db 20401 VFTEPAIAKNPDPPGRCDPPVISNITKDHTVSWKPPADDGSGSPITGYLLEKRETOAVN 20460
Qy 172 ----- 171
Db 20461 WTKVNRKPIIERTLKATGLQEGTEYEFRTAINKAGKPGFSKASKAAYARDPOYPAPP 20520
Qy 172 ----- 171
Db 20521 FPKVYDTRSSVSLSGKPAYDGGSPIIIGYLVEYKRADSDNWVRNCPQLQKTRFEVTG 20580
Qy 172 ----- 178
Db 20581 LMEDTOYQFRVYAVNKIGYSDPSDVPDKHYPKDILIPPEGELDADLRKTLILRAGVTMRL 20640
Qy 179 ----- 178
Db 20641 YVPVGRPPPKITWSKPNVNLDRIGLDIKSTDFDTLRCENYVNDAGKYILTLENSCG 20700
Qy 179 ----- 178
Db 20701 KKEYTIVVKVLDTPGPPVNTVKEISKDSAYVTWEPPIIDGSPPIINVVQKRAERKSW 20760
Qy 179 ----- 178
Db 20761 STVTTECSKTSFRVANLEBCKSYFFRVAENEYIGDGETRDAVKASQTPGPPVDLKV 20820
Qy 179 ----- 178

Db 20821 SVSKSSCSIGMKPHSDGGSRIIGYVVDFTLEENKQWVMKSLSIQISAKDLTEGKEYTF 20880
Qy 179 ----- 178
Db 20881 RVSAENENGECTPSEITVVARDDVVAPDLDLKGLPDLCYLAKENSNFLKIPKIGKPPAPS 20940
Qy 179 ----- 178
Db 20941 VSWKKGEDPLATDTRSVSESSAVNTTLLIVDCQKSDAGKYITILKNVAGTKEGTISIKVV 21000
Qy 179 ----- 178
Db 21001 GKPGIPTPIKFDEVTAEAMTLKWAPPKDGGSEITNYILEKRDVNNKWWTCASAVQKT 21060
Qy 179 ----- 178
Db 21061 TFRVTRLHEGMXYTFRVSAENKYGVGEGLKSEPIVARHPFDPDAPPPNIVDVRHDSVS 21120
Qy 179 ----- 178
Db 21121 LTWTDPKTKGSPITGYHLEPKERNSLWKANKTPIRMRODFKVTGLTEGLEYSFRVMAI 21180
Qy 179 ----- 178
Db 21181 NLAGVGKPSLPSEPVWALDPIIDPPCKPEVINITRNSVTLIWTEPKYDGGHKLGTGIVEKR 21240
Qy 179 ----- 178
Db 21241 DLPSKSMKANHVNVPECAFTVTDLVEGGKYEFRAKANTAGIASPSESTETIICKDEY 21300
Qy 179 ----- 178
Db 21301 EAPTIVLDPIIKDGLTIKAGDTIVLNAISILCKPLPKSSWKAGKDIRPSDITQITSTPT 21360
Qy 179 ----- 178
Db 21361 SSMLTIKYATRKADAGEYITATNPFGTKVEHVKVTVLDVPGPPGVEISNVSAEKATLTW 21420
Qy 179 ----- 178
Db 21421 TPPLDGGSPIKSYILEKRETSRLLTWTVVSEDIQSCRHATKLIQGNIEYIFRSVAVNHYG 21480
Qy 179 ----- 178
Db 21481 KGEVPQSEPVKMDRFGPPGPEKPEVSNVTKNATATVSWKRPVDDGSGSEITGYHVERREK 21540
Qy 179 ----- 178
Db 21541 KSLRWVRAIKTPVSDLRCKVTGLQEGSTYEFVSAENRAGIPPEASDSVLMKDAAYPP 21600
Qy 179 ----- 178
Db 21601 GPPSNPHVTDTKKSASLAWGKPHYDGGLEITGYVVEHQKYDEAWIKDITGTALRITQF 21660
Qy 179 ----- 178
Db 21661 VVPDLOTKKEYNFRIASINDAGCEPAVIPDVEIVEREMAPDFELDAELRRTLVRAGLS 21720
Qy 179 ----- 178
Db 21721 IRIFPVPIKGRPAPEVTWTKDNLNKNRANIENTESFTLLIIECNRYDTGKFMVTIENPA 21780
Qy 179 ----- 178
Db 21781 GKSGFVNVRVLDTPGPVNLNRPTDITKDSVTLHWDLPIDGGSRTITNYIVEKRETRKS 21840
Qy 179 ----- 178
Db 21841 YSTATTKCHKCTYKVTGLSEGCCEYFFRVMAENEYIGEPTEPTEPVRKASEAPSPDLSNI 21900
Qy 179 ----- 178

Db 21901 MDITKSTVSLAWPKPKHGGSKITGYVIEAQRKSDQWTHITTVKGLCEVVRNLTEGEY 21960
QY 179 ----- 178
Db 21961 TFOVMVNSAGRSAPRESRPVIVKEOTMLPELDLRGIYOKLVIAKAGDNKIVEIPVLRP 22020
QY 179 ----- 178
Db 22021 KPTVTWKGDQILKQORVNFETTATSTILNINECVSRSDGPPYPLTARNIVGEVGVII 22080
QY 179 ----- 178
Db 22081 QVHDIPGPTGPIKFDEVSDFVTFWSWDPENDGGVPISNYVVMRQDSTTWELATV 22140
QY 179 ----- 178
Db 22141 IRTTYKATRLTGLEYOFRVKAQRNYGVGPGITSACIVANYPFKVPGPPTQVTAATKD 22200
QY 179 ----- 178
Db 22201 SMTISWHEPLSDGGSPILGYHVERKERNGLMQTVSKALVPGNIFKSSGLTDGIAEYFRV 22260
QY 179 ----- 178
Db 22261 IAEENMAGSKSPKSPMLALDIPDPGKVPPLNITRHTVTLKWAPEYTGFKITSYIV 22320
QY 179 ----- 178
Db 22321 EKRDLPNGRWLKNFNSILENETVSGLTEDAAEYFRVIAKNAAGAI SPPSEPSDAITR 22380
QY 179 ----- 185
Db 22381 DOVEAPKIKVDVKFDTVLKAGAPRLEADVSGRPPTMWSKDGKEGTAKLEIKTA 22440
QY 186 ----- 185
Db 22441 DFSTNLVNDSTRDSGAYTLTATNPGGAKHIFNVKVLDRPGPEGLAVTEVTSEKV 22500
QY 186 ----- 185
Db 22501 LSWFPLDDGGAKIDHYIVOKRETSRLAWTNVASEVQVTKLVTKLLKGNEYIFRVMVN 22560
QY 186 ----- 185
Db 22561 KYGVGEPLESEPVAVNPYGPDPKPNPEVTTITKDSMVVVCWGHPSDGGSEIINIVYER 22620
QY 186 ----- 185
Db 22621 RDKAGORWIKCNKKTLDLRYKVSLTEGHEYEFRIMAENAGISAPSPSPFYKACDTV 22680
QY 186 ----- 185
Db 22681 FRPGPPGNPRVLDTSRSSISIAWNKPIYDGSSEITGYMVEIALPDEDEWQIVTPPAGLKA 22740
QY 186 ----- 185
Db 22741 TSYTITGLTENOYKIRIYAMNSEGLGEPALVPGTPKAEDRMLPPEIELDADLRKVVITR 22800
QY 186 ----- 185
Db 22801 ACCTLRLEVPKGRPAPEVKWARDHGESLDKASIESTSYTLLIVGNVNRFDGKYILTV 22860
QY 186 ----- 185
Db 22861 ENSSGSKSAFNVVRVLDTPGPPQDLKVKEVKTSTVLTDWPPPLDGGSKIKNYIVEKRES 22920
QY 186 ----- 185
Db 22921 TRKAYSTVATNCHKTWKVDQLQEGCSYVFRVLAENEYIGLPAETAESKASERPLPPG 22980
QY 186 ----- 185
Db 22981 KTTLMDVTRNSVLSWEKPEHDDGSRILGYIVEMQTKSGDKWATCATVKVTEATITGLIQ 23040

QY 186 ----- 185
Db 23041 GEEYSFRSAQNEKGLISDPRQLSPVPIAKDLVIPPAFKLLFNTFTVLAGEDLKVDPFFIG 23100
QY 186 ----- 185
Db 23101 RPTPAVTHKDNVPLKQOTTRVNAESTENNLSLLTIKDACREDVGHYVVKLUTNSAGEAIEL 23160
QY 186 ----- 185
Db 23161 NVIVLDKPGPPTGPKVMDEVTDADITLSWGPYPKYDGGSSINNYIVEKRODTSTTWOIVSA 23220
QY 186 ----- 185
Db 23221 TVARTTIKACRLKTGCEYQFRIAANRYCKSTYLNSEPTVAQYPPFKVPGPGTPVVTLS 23280
QY 186 ----- 185
Db 23281 RDSMEVQWNEPISDGGSRVIGYHLERKERNLSILWVLNKTPIPTQTKFKTTGLEEGVEYEF 23340
QY 186 ----- 185
Db 23341 RVSAENIVGIGKPSKVSECYVARDPCDPPRPEAIIVTRNSVTLOWKKPTYDGGSKITCY 23400
QY 186 ----- 185
Db 23401 IVEKKELPEGRWMAKASFTNIIDTFEVTGLVEDHRYEYFRVIAARNAAGVESESESTGAT 23460
QY 186 ----- 185
Db 23461 ARDEVDPPIRISMDPKYKDTIVVHAGESFKVDADIYKPIPTIQWIKGDOELSNATARLEIK 23520
QY 186 ----- 185
Db 23521 STDFATSLSVKDAVRVDSGNVILKAKNAVAGERSVTNVKVLDRPPEGPVVISGVTAEK 23580
QY 186 ----- 185
Db 23581 CTLANKPPLQDGGSDIINIVYERRETSLRVTVVDANVQTLSCVKTKLLEGNEYTRINA 23640
QY 186 ----- 185
Db 23641 VNKYGVGEPLESEPVVAKNPFVVPDAPKAPEVTTVTKDSMIVVWBERPASDGGSEILGYVL 23700
QY 186 ----- 185
Db 23701 EKRDKEGIRWTRCHRRLLIGELRLRVTLGLENHDYEFVSAENAGLSEPPSPSAYOKACD 23760
QY 186 ----- 185
Db 23761 PIYKPGPPNNPKVIDITRSSFVLSWSKPIYDGCCEIQGYIVEKCDVSGEWTMCTPTTCI 23820
QY 186 ----- 185
Db 23821 NKTNIEVEKLEKHEYNFRICAINKAGVHADVPGPPIVEKLEAPDIDLDELRLKIIN 23880
QY 186 ----- 185
Db 23881 TRAGSLRFLFVPKGRPTEPVKWGVKVDGEIRDAAIIDVTSSFTSLVLDNVNRYDSGKYTL 23940
QY 186 ----- 185
Db 23941 TLENSSGTKSAFVTVRVLDTPSPPNVNLKVTETIKDSVSTIWEPPPLDGGSKIKNYIVEK 24000
QY 186 ----- 185
Db 24001 EATRKSAAVVTNCHKNSWKIDQLQEGCSYVFRVTAENEYIGLPAQTADPIKVAEVP 24060
QY 186 ----- 185
Db 24061 PKITVDVTRNSVLSWTKPEHDDGSKIIQYIVEMQAKHSEKSECAVKSQAQVITNL 24120

QY 186 ----- 185
Db 24121 TOGEEYLFRRVAVYNEKGRSDPRSLAVPIYAKDLVIEPDVKPAPSSYSVQVQDLKIEVPI 24180
QY 186 ----- 185
Db 24181 SGRPKFTITWKDGLPLKOTTRINVTDSLDTLSIKETHKDDGGQYGITVANVVGKTA 24240
QY 186 ----- 185
Db 24241 SIEIVLDRPKPPKGVKFDVDSAESITLSWNPPLYTGGCQITNYIVQKRDTTTTVDVV 24300
QY 186 ----- 185
Db 24301 SATVARTLKVKLTGTGYQFRIFAENRYGQSFALSDPIVAQYPYKBPPTPFATA 24360
QY 186 ----- 185
Db 24361 ISKDSMVIOMHEPVNNGSPVIGYHLERKERNISILWTKVNTLIHTQFKAQNLBEGIEY 24420
QY 186 ----- 185
Db 24421 EFRVYAENIVGKASKNSECYVARDPCDPPGTPPEIMVKRNEITLQWTKPVYDGGSMIT 24480
QY 186 ----- 185
Db 24481 GYIVEKRDLPDGRMKASFNTVETQFTVSGLTEDQRYEFRVIKNAAGAIKSPSDTGP 24540
QY 186 ----- 185
Db 24541 ITAKDEVLPRIKSMDFKRTIIVNAGETFRLEADVHGKPLPTIEWLRGDKETEESARCE 24600
QY 186 ----- 185
Db 24601 IKNTDFKALLIVKDAIRIDGGVILRASNVAGSKSPVNVKVLDRPGPEGPVQVTGVS 24660
QY 186 ----- 185
Db 24661 EKCSLTWSPLODGGSDISHYVVEKRETSRLAWTVVASEVVTNSLKVTKLLEGNEYVFI 24720
QY 186 ----- 185
Db 24721 MAVNKYGVGEPLSAPVLMKNPFLPGPKPLEVTNIAKDSMTVCWNRPSDGGSEIIGY 24780
QY 186 ----- 185
Db 24781 IVEKDRSGIRWIKCNKRITDLRLRVTLGLTEDHEYEFRVSAENAAAGVGEPSATVYYKA 24840
QY 186 ----- 185
Db 24841 CDPVFKPGPTNAHIVDTTKNSITLAWGKPIYDGGSEILGYVEICKADEEWOIVTPQT 24900
QY 186 ----- 185
Db 24901 GLRVTRFEISKLEHOEYKIRVCALNKVGLGEATSVPGTVKPEDKLEAPELDLDELKRG 24960
QY 186 ----- 185
Db 24961 IIVRAGSARIHIPFKGRPTPEITWSREGEFTDKVQIEKGVNYTOLSIDNCDNRDAGKY 25020
QY 186 ----- 188
Db 25021 ILKLENSGSKSAFTVTKVLDTPGPPQNLAVKEVRKDSAFVWEPPIIDGAKVKNYVID 25080
QY 189 ----- 188
Db 25081 KRESTRKAYANSSKSKTSFKVENLTEGAIYYFRVMAENEFVGVGVVETVDAAKAEPP 25140
QY 189 ----- 188
Db 25141 SPPGKVTLTDVTSASLWKEPHEHGGSRVLGYVVEMQPKGTEKWSIVAESKVCNAVVT 25200
QY 189 ----- 188

Db 25201 GLSSGOEQYQFRVKAYNEKGKSDPRVLGVPVIAKDLTIQPSLKLFPENTYSIQAGEDLKIEI 25260
QY 189 ----- 188
Db 25261 PVIGRPRPNISWYKDGELPKQTTTRVNVEETATSTVLHIKEGNKDDFCYKTVTATNSAGTA 25320
QY 189 ----- 188
Db 25321 TENLSVIVLEKPGPPVCPVRFDEVDSADFVVVISWEPPPAYTGGCQISNVIIVEKRDTTTTTWH 25380
QY 189 ----- 188
Db 25381 MVSATVARTTIKITKLKTGTGYQFRIFAENRYGKSAPLDSKAVIVQYFPKPGPGPTPFV 25440
QY 189 ----- 188
Db 25441 TSISKDQMLVQWHEPVNDGSKTIIGYHLEQEKNSILWVKNLKTPIQDTKFKTTGLDEGL 25500
QY 189 ----- 188
Db 25501 EYEFKVSANIVGIGKPSKVSECFVARDPCDPPCRPEAIVITRNNTLWKKPAYDGGSK 25560
QY 189 ----- 188
Db 25561 ITGYIIVEKKDLPDGRMKASFNTVLETFVSGLVEDQRYEFRVIARNAAGNFSEPSDSS 25620
QY 189 ----- 188
Db 25621 GAITARDEIDAPNASLDPKYKDVIVVHAGETFVLEADIRGKPIPDVWVMSKDGKELEETA 25680
QY 189 ----- 188
Db 25681 RMEIKSTIOKTTLVVVKDCIRTDGGQYILKLSNVGGTKSPITVKVLDLRPGPPGPKLVTC 25740
QY 189 ----- 193
Db 25741 VTAEKYLAWNPPLODGGANISHYIIEKRETSRLSWTQVSTEVOALNKYVTKLLPGNEYI 25800
QY 194 ----- 193
Db 25801 FRVMAVKNYKIGEPLESGPVACNPYKPPGPPSTPEVSATIKDSMVVTVARPVDDGGTEI 25860
QY 194 ----- 193
Db 25861 EGYILEKRDKEGYRWTCKNKKTLTDLRLRVTLGLTEGHSYEFVRAAENAAAGVGEPSPSVF 25920
QY 194 ----- 193
Db 25921 YRACDALYPPGPPSPNKVTDTSRSSVSLAWSKPIYDGGAPVKGYVVEVKEAAADEWTTCT 25980
QY 194 ----- 193
Db 25981 PPTGLOGKQFTVTKLKENTYENFRICAINSEGVGEPATLPGSVVAQERTPEPPELDADL 26040
QY 194 ----- 193
Db 26041 RKVVVLRASATLRLFVTIKGRPEPEVKWEKAEGLTDRAGIEVTSSTFMLVIDNVTFRDS 26100
QY 194 ----- 193
Db 26101 GRYNLTLENNSGSKTAFVNVVRLDPSAPVNLTIREVKKDSVTLISWEPPLIDGGAKITNY 26160
QY 194 ----- 193
Db 26161 IVEKRETRKAYATITNNCTKTTFRLENLOEGCSYFRVLASNEYGIGLPAETTEPVKVS 26220
QY 194 ----- 193
Db 26221 EPPLPGRVTLVDVTRNTATIKWEKPESDGGSKITGVVEMQTKGSEKWSCTQVKTLEA 26280
QY 194 ----- 193

Db 26281 TISGLTAGEEYFRVAANNEKGRSDPRQLGVPIARDIEIKPSVLEPHTFNVKAREQLK 26340
Qy 194 ----- 193
Db 26341 IDVPEKGRQATVNRKDGQTLKETTRVNSSKIVTSLIKEASKEDVGIELCVNSNA 26400
Qy 194 ----- 193
Db 26401 GSITVPIIIVLDRPGPIRIDEVSCDSITISWNPEYDGGCOISNVIVEKETTSTT 26460
Qy 194 ----- 193
Db 26461 WHIVSOAVARTSIKIVRLTTGSEYQFRVCAENRYGKSSYESSAVVAEYFPFPGPGPT 26520
Qy 194 ----- 193
Db 26521 KVVHATKMTLWQVPVNDGSGSRVIGYHLEYKERSILWSKANKILLIADTOMKVSGLDE 26580
Qy 194 ----- 193
Db 26581 GLMEYRVVAENIAGIGKSKCEVPARDPCDPGQPEVTNITRKSYSKLWSKPHYDGG 26640
Qy 194 ----- 193
Db 26641 AKITGIYVERRELDPGRMLKCNITNIQETIYFEVTELTEDQRYEYFRVAFARNADSVSEPE 26700
Qy 194 ----- 193
Db 26701 STGPIIVKDDVEPRVMDKFRDVIIVKAGEVLKINADIAGRPLPVISWAKDGEIEER 26760
Qy 194 ----- 193
Db 26761 ARTEIISTDNHTLLTVKDCIRRDGTQYVLTILKNVAGTRSAVAVCKVLDKPGPPAGPLEIN 26820
Qy 194 ----- 193
Db 26821 GLTAEKCSLSWGRPOEDGGADIDYIYVEKRETSHLAWTICEGELQMTCKVTLLKGNEY 26880
Qy 194 ----- 193
Db 26881 IFRVTGVNKYVGGEPLESVAIKALDPFTVPSPPTSLEITSVTKESMTLCWSRPSDGGSE 26940
Qy 194 ----- 193
Db 26941 ISGYIIRERKSLRWVRNKKPVYDLRVKSTGLREGCEYEVRYVAENAAGLSLPSETSP 27000
Qy 194 ----- 193
Db 27001 LIRAEADPVLSPSPSKPIVDGKTTITIAWVKPLFDGGAPITGYTVEYKKSDDTDWKTS 27060
Qy 194 ----- 193
Db 27061 IQSLRGTEYITISGLTGAEYFRVKSVMKVGASDPSDSDPQIAKEREELPFDIDSEMR 27120
Qy 194 ----- 193
Db 27121 KTLIVKAGASFTMTVPFRGRPPVNLWSKPDOLTRAYVDTTDSRTSLTIENANRNDG 27180
Qy 194 ----- 193
Db 27181 KYTLTIQNLVSAASLTLVVKVLDTPGPPTNTITVQDVTKEASAVLSWDVPENDGGAPVKNYH 27240
Qy 194 ----- 193
Db 27241 IEKREASKAWSVTNNCNRLSYKTNLQEGAIYFRVSGENEPGVGIPAEKGVKITE 27300
Qy 194 ----- 193
Db 27301 KPSPEKLGVTISKSDSVSLTWLKEPHDGGSRIVHYVVEALEKGKNWVKCAVAKSTHHV 27360
Qy 194 ----- 193
Db 27361 VSGLRSENSEYFRFAENQAGLSOPRELLLPVLVLIKEQLEPPEIDMKNFPSTHYVVRAGSN 27420

Qy 194 ----- 193
Db 27421 LKVIDIPISGKPLPKVTLSDRGVPLKATMRFNTEITAENLTINLKESVTADAGRYEITAA 27480
Qy 194 ----- 193
Db 27481 SSGTTKAFINIVLDRPGPTGPVVISDITEBSVTLKWEPPKYDGGSOVTNYILLKRETS 27540
Qy 194 ----- 193
Db 27541 TAVMTVSATVARTMMKVNKLTTGEEYQFRIKAENRFGISDHIDSACVTVKLPYTPGPP 27600
Qy 194 ----- 193
Db 27601 STPWTVNTRESITVGVHEPVNGSGSAVVGYHLEMKDRNSILWQKANKLVIRTHFKVTT 27660
Qy 194 ----- 193
Db 27661 ISAGLIYEFYVAENACGVGKPSHPSEPVLAIDACEPPRNVRIITDISKNSVLSLWQOPAF 27720
Qy 194 ----- 193
Db 27721 DGSKITGYIVERDLDPGRWTKASFNTVETQFIISGLTQNSQYEFYFARNAVGSIN 27780
Qy 194 ----- 193
Db 27781 PSEVVGPICTIDSYGGPVLDLPLEYTEVVKYRAGTSVKLRAGISGKPAPTIEWYKDKEL 27840
Qy 194 ----- 193
Db 27841 QTNALVCVENTDOLASILIKDADRNLSCYELKLRNANGSASATIRVOILDKPGPPGGPI 27900
Qy 194 ----- 193
Db 27901 EFKVTAEKITLLWRPPADGGAKITHYIYVEKRETSRVVSMVSEHLEECITTTTKIIG 27960
Qy 194 ----- 193
Db 27961 NEYIFRVAVNKYGIGEPLESDSVVAKNAFVTPGPPGIPEVTKITKNSMTVWSRPIADG 28020
Qy 194 ----- 193
Db 28021 GSDISGYFLEKRDKSLGWFKVLKETIRDRQKVTGLTENSQYVRVCAVNAAGQGPSE 28080
Qy 194 ----- 193
Db 28081 PSEFYKAADPIDPPGPPAKIRIADSTKSSITLWWSKPVYDGSQAVTYVVEIROGEEEW 28140
Qy 194 ----- 193
Db 28141 TTVSTKGEVRYTEYVVSNLKPGVNYIFRVSANCAOGGEPTEMNEPVQAKDILEAPEIDL 28200
Qy 194 ----- 193
Db 28201 DVALRTSVIAKAGEDVQVLIPFKGRPPETVTRKDEKNLGSDARYSTIENTDSSLLTIPQ 28260
Qy 194 ----- 193
Db 28261 VTRNDTGKYLTIENGVEPKSSTSVKVLDTPAACQKLOVKHVSRTVTLMLDPPPLIDG 28320
Qy 194 ----- 193
Db 28321 GSPINIYVIEKRDATKRTWSVVSHKCSSTSPKLDLSEKTPFFFRVLAENEIGIGECET 28380
Qy 194 ----- 193
Db 28381 TEPVKAEEVPAPIRDLMSKOSTKTSVILSWTKPDFDGGSVITEYVVERKKGEOQWTSAG 28440
Qy 194 ----- 193
Db 28441 ISKCTEIVSOLKEQSVLEFRVFAKNEKGLSDPVTIGITVTKELIITPEVDLSIDPGAQV 28500

Qy 194 ----- 193
Db 28501 TVRIGHNVHLEPYKGRPKPSISMLKDLPLKESEFFVRFSTENKIITLSIKNAKEHGK 28560
Qy 194 ----- 193
Db 28561 YTVILDNAVCRIAPITVITLGPSPKPGPIRFDEIKADSILSWDPEDNGGGEITCYS 28620
Qy 194 ----- 193
Db 28621 IEKRETSOTNMKVCSSVARTTFKVPNLVKADEYQFRVRAENRYGVSOPLVSSIIVAKHQ 28680
Qy 194 ----- 193
Db 28681 FRIPGPGKPYIYNVTSOGMSLTWDAPVYDGGSEVTFGHVEKKERNSTLWKVNTSPISG 28740
Qy 194 ----- 193
Db 28741 REYRATGLVGLDYQFRVYAENSAGLSSPSPSKFTLAVSPVDPGTPDYIDVTRTITL 28800
Qy 194 ----- 193
Db 28801 KWNPLRDGSGKIVGYSTIEKKOGRNWRVRCNFTDVSECQYTVTGLSPGDRYEFRIIARNA 28860
Qy 194 ----- 196
Db 28861 VGTISPPSSQSGIIMTRDENVPPIVEFGPEYFDGLIIKSGESLRIKALVQGRPVPRVTFW 28920
Qy 197 ----- 196
Db 28921 KDGVEIKRMNMEITDVLGSTLFVRDATRDHRGVYTVVEAKNAGSAAEKVKVQDTPG 28980
Qy 197 ----- 196
Db 28981 KVGPIRFTNITGEKMTLWMDAPLNDGCAPITHYIIIEKRETSRLAWALIEDKCEAQSYTA 29040
Qy 197 ----- 200
Db 29041 IKLINGNEYQFRSAVNFVGRPLDSDPVVAQIQYTVDPAGPIEPSNITGNSITLWA 29100
Qy 201 ----- 200
Db 29101 RPESDGGSEIOQYILERBEKKSTRWKVVISKRPISETRFKVTGLTEGNEYEFHVMAENAA 29160
Qy 201 ----- 200
Db 29161 GVGPASGISRLIKCREPVNPPGPTTVVKVTDTSKTTVSLENSKPVFDGGMELIIGYIEMC 29220
Qy 201 ----- 200
Db 29221 KADLGDNHKNVNAECVKTRYTVDLQAGEEYKFRVSAINGAGKGDSCVETGTIKAVDRLT 29280
Qy 201 ----- 200
Db 29281 APELDIDANFKQTHVVRAGASIRLFYAYQGRPTPTAVWSKPDNSLRLADITHDTSFSTL 29340
Qy 201 ----- 200
Db 29341 TVENCNRNDAGKYTLTVENNSGSKSITFTVKVLDTPGPPGPIITFKDVTGRSATLWMDAPL 29400
Qy 201 ----- 200
Db 29401 LDGGARIHHYVVEKREASRRSQVISEKTRQIFKVNDAEGVPYFRVSAVNEYGVGEP 29460
Qy 201 ----- 200
Db 29461 YEMPEPIVATEQAPRRRLDVVDTSKSSAVLAWLKPDDHGGSRITGYLLEMRQKGSDFW 29520
Qy 201 ----- 200
Db 29521 EAGHTKQLTFTVERLVEKTEYFRVKAKNADAGYSEPREAFSSVIIKEPOIETADLTGIT 29580
Qy 201 ----- 200

Db 29581 NQLITCKAGSPFTIDVPISGRPAKVTWKLEMLKETDRYSITTTKORTTLTVKDSMRG 29640
Qy 201 --GSFFL----- 205
| : ||
Db 29641 DSGRYFLTLENTAGVKTFSVTVVVIGRPGVPTGPIEVSSVSAESCVLWSGPKDGGGTEI 29700
Qy 206 ----- 205
Db 29701 TNYIVEKRESCTTAWOLVNSSVKRTOIKVTHLTMYEYSFRVSSNRFVSKPLESAPII 29760
Qy 206 ----- 205
Db 29761 AEHPFVPSPAPTRPEVYHVSANAMSIRWEEPYHDGGSKILGYWVEKKERNITLWKENKV 29820
Qy 206 ----- 205
Db 29821 PCLECNKYKVTGLVEGLEQYQFRYALNAAGYSKASEASRPIMAQNPVDAPGRPEVTDVTRS 29880
Qy 206 ----- 205
Db 29881 TVSLWSNAPYDGGSKVVGYIIERKPVSEVGDGRWLKCNYYTIVSDNFFVTALSEGDTYE 29940
Qy 206 ----- 205
Db 29941 FRVLAKNAAGVISKSESTGPTCRDEYAPPKAELDARLHGLDVTIRAGSDLVLDAAVGG 30000
Qy 206 ----- 205
Db 30001 KPEPKIIWTKGDKELDCEKVSLOYTGKRATAVIKFCDRSDSGKYTLTVKNASGTFKAVSV 30060
Qy 206 ----- 205
Db 30061 MKVLDSPGCGKLTVSRTQEKCTLAWSLPQEDGGAETHYIVERRETSRLNWNVIVEGE 30120
Qy 206 ----- 205
Db 30121 CPTLSYVVTRLIKNNEYIFRVRANVKGPGVVESEPIVARNSTFIPSPGPIEVEGTGK 30180
Qy 206 ----- 205
Db 30181 EHIIQWTKPESDGGNEISNLYLDKREKKSRLWRVNKYVYVDTRLKVTSLMEGCDYOF 30240
Qy 206 ----- 205
Db 30241 RVTAVNAGNSEPSEASNFISCREPSYTPGPPSAPRVYDTTKHSISLAWTKPMYDGGTDI 30300
Qy 206 ----- 205
Db 30301 VGYVLEMQEKTDQWYRVHTNATIRNTEFTVPDLKMGQKYSFRVAANVKGMYSEYSESIA 30360
Qy 206 ----- 205
Db 30361 EIEPVERIEIPDELAADDLKKTTVIRAGASLRLMVSVSRPPPVITWSKQIDLASRAII 30420
Qy 206 ----YSKLTVDK----- 213
|||
Db 30421 DTTESYLLIYDKNRYDAGKYTIEAENQSGKKSATVLVKYDTPGPCPSVKVKEVS RDS 30480
Qy 214 ----- 213
Db 30481 VTITWEIPTIDGAPVNNYIVEKREAAAMRAFKTVTTKCKTLYRISGLVEGTMYYFRVLP 30540
Qy 214 ----- 213
Db 30541 ENIYGEPCEPETSDAVLVSEVPLVPAKLEVVDVTKSTVTLAWEKPLYDGGSRLTGYVLEA 30600
Qy 214 ----- 213
Db 30601 CKAGTERMKVVTLKPTVLEHTVTSLNEGEQYLFRIRAOKEGVSEPRETVTAVTVQDLR 30660
Qy 214 ----- 213

Db 30661 VLPTIDLSTMPQKTIHVPAGRPVELVPIAGRPPPAASWFFAGSKLRESERVTVETHKV 30720
QY 214 ----- 213
Db 30721 AKLTIRETTIRDTGEYTLKKNVTGTTSETIKVILDKPGPPGPIKIDEIDATSTISW 30780
QY 214 ----- 213
Db 30781 EPELDGGAPLSGVVEQDAHRPGWLPVSESVTRSTFKFTRLTEGNEYVFRVAATNRFQ 30840
QY 214 ----- 213
Db 30841 IGSYLQSEVIECRSIRIPGPETLQIFDVSRDGMTLTWYPEDDGGSQVTGYIVERKEV 30900
QY 214 ----- 213
Db 30901 RADRWVRNVKVPVMTYRSTGLTEGLEYEHRVTAINARGSGKPSRPSKPIVAMDPIAPP 30960
QY 214 ----- 213
Db 30961 GKQNPRTVDTTRTSVSLAWSVPEDGGSKVTGYLIEMQKVDQHEWTKCNTTPTKIREYT 31020
QY 214 -----SRMOQG----- 219
|:|:
Db 31021 LTHLPOGAERYFRVLACNAGGPGPAEVPGVTKVTEMLEYDPYELDERYOEGIFVRQGV 31080
QY 220 ----- 219
Db 31081 IRLTIPIKGPPICKWKKEQDISKRAMIATSETHTELVIKADRGDSGYDLVLENKC 31140
QY 220 ----- 219
Db 31141 GKAVYIKVRVIGSPNSPEGPLEYDDIQVRSVRVSNRPADGGADILGILYLERREVPKA 31200
QY 220 ----- 219
Db 31201 AWYTIDSRVGRGTSLVVKGLKENVEYHFRVSAENQFGISKPLKSEPEVTPKTLNPPPEPS 31260
QY 220 ----- 219
Db 31261 NPPEVLDTKSSVLSWSRPKDDGSRVGYIYIERKETSTDKWVRHNTQITTTMTYTVTG 31320
QY 220 ----- 219
Db 31321 LVPDAEQFRIIAQNDVGLSETSPASEPVVCKDPFDKPSQPGLEILSISKDSVTLQWEK 31380
QY 220 ----- 219
Db 31381 PECDDGKEILGYWVEYRQSGDSAMKSNKERIKDKOFTTIGLLEATEYEFVRFAENETGL 31440
QY 220 ----- 219
Db 31441 SRPRRTAMSIKTLTSGEAPGIRKEMKDVTTKLGEAAQLSCQIVGRPLPDIKWYRFGKEL 31500
QY 220 ----- 219
Db 31501 IOSRKYKMSSDGRHTLTVMTEBEQDEGVYTCIATNEVGEVETSSKLLQATPQHPGPYP 31560
QY 220 ----- 219
Db 31561 LKEKYYGAVGSTLRLHVMYIGRPVPAMTWFGHQQKLLONSENITIENTEHTHLVMKNVOR 31620
QY 220 -----NVF----- 222
|:|:
Db 31621 KTHAGYKVLNSVFGTVDAILDVEIQDKPDKPTGTGPIVIEALLKNSAVISWKPADDGGS 31680
QY 223 ----- 222
Db 31681 WITNYVVERCEAKEGAEMOLVSSAISVTTCRIVNLNTENAGYVFRVSAQNTFGISDPLEVS 31740
QY 223 ----- 222
Db 31741 SVVLIKSPEKPGAPGKPTITAVTKDSCVVAWKPASDGGAKIRNYYLEKREKKONKWIS 31800

QY 223 ----- 222
Db 31801 VTTEEIRETVFSVKNLIEGLEVEFRVKCNLGGSESEISEIPITPKSDVPIQAPHPKEE 31860
QY 223 ----- 222
Db 31861 LRNLNVYQSNATLVCKVTGHPKPIVKWYRQGEIADGLKYRIQEFKGYHQLIASVT 31920
QY 223 ----- 222
Db 31921 DDDATVYQVRATNOGGSVSGTASLEVEVPAKIHLPKTLEGMAVHALGGEVWSIKIPFSG 31980
QY 223 ----- 222
Db 31981 KPDPVITWQGGQDLIDNNGHYQVIVTRSTSLVFPNGVERKIDAGFYVVCANKNRFIDQKT 32040
QY 223 ----- 222
Db 32041 VELDVADVPDPRGVKVSVDVSRDSVNLWTPEPASDGSKITNYIVEKCATTAERWLRVGQ 32100
QY 223 ----- 222
Db 32101 ARETRYVINLFCKTSYQFRVIAENKFGLSKPSSEPTITTKEDTRAMNYDEEVDETRE 32160
QY 223 ----- 222
Db 32161 VSMTKASHSTKELYEKYMAIEDLGRGEGIVHRCVETSSKKTYMAKFKVKGTDQVLVK 32220
QY 223 ----- 222
Db 32221 KEISILNIAHRNHLHSHESFESMEELVMFEFISGLDIFERINTSAFELNEREIVSYVH 32280
QY 223 ----- 222
Db 32281 QVCEALQFLHSHNIGHDIRPENIIYQTRRSSTIKIIEFGARQLKPGDNFRLLFTAPEY 32340
QY 223 ----- 222
Db 32341 YAPEVHQHDVSTATDMWSLGLTVLVLLSGINPFLAETNQIENIMNAEYTFDEEAPKE 32400
QY 223 ----- 222
Db 32401 ISIEAMDFVDRLLVKERKSRMTASEALQHPWLKQKIERVSTKVIRTLKHRRYYHTLLKKD 32460
QY 223 ----- 222
Db 32461 LNMVYSAARISCGGAIRSQKGSVAKKVASIEIGPVSQIMHAVGEGGHVYVCKIEN 32520
QY 223 ----- 222
Db 32521 YDOSTQVTVYFGRQLENSEKYEITYEDGVALIYVKDITKLDGTYRCKVWVNDYGEDSSY 32580
QY 223 ----- 222
Db 32581 AELFKGVREVYDYICRRTMKKIKRRDTMRLLERPPEFTLPYKNTAYGVNRFVGTI 32640
QY 223 ----- 222
Db 32641 TVHPEPHVTWYKSGOKIKPGDNDKKYTFESDKGLYQLFTINSVTTDDAEYTVVARNKYGE 32700
QY 223 -SC-|||----- 224
Db 32701 DSCAKLAVTLHPPTDSTLRPMFKRLLANAEQSQSVCFEIRVSGIPPPPTLKWEKDCQ 32760
QY 225 -----SVMHEAL-:|:|----- 231
Db 32761 PLSLGPNIETIIHEGLDYVALHIRTLPEDTYGYRVATATAGTSCQAHQLQVERLRYKKQ 32820
QY 232 ----- 231
Db 32821 EFKSKEEHERHVQKOIDKTLRMAEILSGTESVPLTQVAKEALREAAVLYKPAVSTKTVKG 32880

```
QY 232 ----- 231
Db 32881 EFRLEIEEKERKLMPYDVPEPRKYQFTIEDQRIKQFVPMQDMKWYKIRDOYEMP 32940
QY 232 ----- 231
Db 32941 GKLDVVQKPKRIRLSRWQFYVMPPLRITDQYRKPWRIPKLSQDDLEIVRPARRTPS 33000
QY 232 ----- 231
Db 33001 PDYDFYRPRRSLGDISDELLPIDDYLAMKRTBEERLRLBEELELGSASPPSRPP 33060
QY 232 ----- 231
Db 33061 HFEISSLRYSSQAHVKVEETRKDFRSTVHIPTKAEASTSYAELRERHAQAAYRQPKOR 33120
QY 232 ----- 231
Db 33121 ORIMAEREDELLRPVTTTQHLSEYKSELDPMSEKSRKSRQREVTEITEIEEYEI 33180
QY 232 ----- 231
Db 33181 SKHAQRESSSSASLLRRRLSFTYIELMRPVSELIRSPQPAEYEDDTERSPPTPER 33240
QY 232 ----- 231
Db 33241 TRPRSPVSPSSRSRFSARFDIFSRYESMKAALKTKTQTSERKVEVLSQPFLLDHA 33300
QY 232 ----- 233
Db 33301 PRITLRMRSHRVPQONTREFILNVQSPKTAEVKWHNGVELQSSKIHNTSGVLTLEI 33360
QY 234 ----- 233
Db 33361 LCHTDGTYRAVCTNYKGEASDYATLDVTGGDYTYASQRDEEVPRSVFFELRTEA 33420
QY 234 ----- 233
Db 33421 YAVSSFKKTSMEBASSREVSKQMTETRESLSYHSASAEKSALEEKSLKSTTR 33480
QY 234 ----- 233
Db 33481 KIKTTLAARILTKPRSMTVYEGESARFSCDTGEPVTVTLRKQGLVTSARHQTITK 33540
QY 234 ----- 233
Db 33541 YKSTFEISSVQASDEGNYSVVENSEGKQAEFTLTIQKARVTEKAVTSPPRVKSPEPRV 33600
QY 234 ----- 233
Db 33601 KSPEAVKSPKRVKSPSPKAVSPTETKPTPEKVQHLVPSAPPKITQFLKAEASKEIA 33660
QY 234 ----- 235
Db 33661 KLTCWESSVLRAKEVTWYKDGKKLKGHGFQFHSADGTGYELKINNLTSDQGEYVCEI 33720
QY 236 ----- 235
Db 33721 SGEIGTSKTNLQPMQAFKSIHEKSVKISETKKSDQKTTESTVTRKTEPKAPEISSKP 33780
QY 236 ----- 235
Db 33781 IVTGLQDVTYSSDVAKFAVKATGEPRPTAIWKDGKAITQGGYKLSDEKGGFFLEIHK 33840
QY 236 -----
Db 33841 TDTSDGLYCTVKNKSAGSVSSCKLTIKAIKDTQAKQVSTQKTSEITTPQK 33891
|||: :||
TQKSLSLSPK 246
RESULT 14
Q10466
ID Q10466 PRELIMINARY; PRT: 26926 AA.
AC Q10466;
```

```
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Titin, heart isoform N2-B (EC 2.7.1.-) (Connectin).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HEART;
RX MEDLINE=96026330; PubMed=7569978;
RA Labeit S.; Kolmer B.;
RT "Titins: giant proteins in charge of muscle ultrastructure and
RT elasticity";
RL Science 270:293-296(1995).
RN [2]
RP SEQUENCE OF 2277-25376 FROM N.A.
RX MEDLINE=92258380; PubMed=1582406;
RA Labeit S.; Gautel M.; Lakey A.; Trinick J.;
RT "Towards a molecular understanding of titin.";
RL EMBO J. 11:1711-1716(1992).
RN [3]
RP SEQUENCE OF 1976-2014 FROM N.A.
RA Labeit S.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=95331314; PubMed=7607248;
RA Gautel M.; Castiglione-Morelli M.A.; Pfuhl M.; Motta A.; Pastore A.;
RT "A calmodulin-binding sequence in the C-terminus of human cardiac
RT titin kinase.";
RL Eur. J. Biochem. 230:752-759(1995).
CC -!- FUNCTION: THIS GIANT MUSCLE PROTEIN MAY BE INVOLVED IN MUSCLE
CC ASSEMBLY AND IN MAINTAINING THE STRUCTURAL INTEGRITY OF
CC SARCOMERES. MAY HAVE PROTEIN KINASE ACTIVITY.
CC -!- ALTERNATIVE PRODUCTS: A NUMBER OF FORMS OF THIS PROTEIN ARE
CC PRODUCED BY ALTERNATIVE SPLICING WHICH DIFFER IN TISSUE
CC DISTRIBUTION. DIFFERENT SIZE TRANSCRIPTS MAY ALSO EXIST WITHIN ANY
CC ONE TISSUE. THE SEQUENCE SHOWN HERE IS THAT OF THE HEART ISOFORM
CC N2-B.
CC -!- TISSUE SPECIFICITY: MUSCLE-SPECIFIC.
CC -!- SIMILARITY: TO THE CATALYTIC DOMAINS OF OTHER SERINE/THREONINE
CC KINASES.
CC -!- SIMILARITY: BELONGS TO IMMUNOGLOBULIN SUPERFAMILY. CONTAINS 112
CC IMMUNOGLOBULIN C2-LIKE DOMAINS AND 132 FIBRONECTIN TYPE III-LIKE
CC DOMAINS.
DR EMBL; X64698; CAA45939.1; -
DR EMBL; X83270; CAA58243.1; -
DR EMBL; X64697; CAA45938.1; -
DR EMBL; X90568; CAA62188.1; -
DR EMBL; X64699; CAA45940.1; -
DR HSSP; P56276; ITLK.
DR InterPro; IPR000282; Cytok_receptor_2.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR000577; FGGV_kin.
DR InterPro; IPR003962; FNIII_repeat.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR000129; Peptidase_S24.
DR InterPro; IPR002016; Peroxidase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00041; fn3; 132.
DR Pfam; PF00047; ig; 91.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00014; FNTYPEIII.
DR PRINTS; PR00726; LEXASERPTASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00060; FN3; 127.
```

DR SMART; SM00408; IGC2; 23.
DR SMART; SM00410; IG_like; 79.
DR SMART; SM00220; S_TKG; 1.
DR PROSITE; PS00933; FGGY_KINASES_1; UNKNOWN_1.
DR PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW Muscle protein; Cytoskeleton; Structural protein; Calmodulin-binding;
KW Serine/threonine-protein kinase; Alternative splicing; Repeat;
KW Immunoglobulin domain; Phosphorylation.
FT DOMAIN 1370 1389
FT DOMAIN 429 4614
FT DOMAIN 24731 25070
FT DOMAIN 25030 25056
FT MOD_RES 1372 1372
FT MOD_RES 1377 1377
FT MOD_RES 1382 1382
FT MOD_RES 1387 1387
FT MOD_RES 26171 26171
FT MOD_RES 26178 26178
FT MOD_RES 26184 26184
FT MOD_RES 26190 26190
FT CONFLICT 22277 22277
FT CONFLICT 22449 22449
FT CONFLICT 22454 22454
FT CONFLICT 23324 23324
SQ SEQUENCE 26926 AA; 2993428 MW; DSEECD3254DF5523 CRC64;
Query Match 34.7%; Score 463; DB 4; Length 26926;
Best Local Similarity 0.7%; Pred. No. 0.00021;
Matches 168; Conservative 43; Mismatches 34; Indels 23731; Caps 53;
QY 2 WLK----- 4
DB 1543 WLKNSDIIVPHKPIRIEGTGEALAKIDSTVSQDSAWYTATANKAGRDTRCKVNV 1602
QY 5 -----AFYDK----- 9
DB 1603 VFAPEPEPERKLIIPRGTYRAKEIAAPELEPLHLRYGQEWEEGLDYDKERQKPPFKK 1662
QY 10 ----- 9
DB 1663 LFSRLKRFPAHFECRLTPISDPTMVVWVLDHGRPLEAANRLRINEFGYCSLDYGVAY 1722
QY 10 ----- 9
DB 1723 SRDSGIITCRATNKGTDTHTSATLIVKDEKSILVERSQLPEGRKGLQRIEELERMAHEGAL 1782
QY 10 ----- 9
DB 1783 TGVTTDQKEQKPDIVLYPEPVRVLEGTARFCRCVTGYPPQKVNWVNLGQLIRKSRFR 1842
QY 10 -----VAE----- 12
DB 1843 VRYDGIHLDIVDCKSYDTGEVKVTAENPEGVIEHKVKLEIQQRDEFRSLRRAPEPRPE 1902
QY 13 -----KLKEA----- 17
DB 1903 FVHPEPGKQFVQKVRPVDTTETKEVVKLRAERTHEKVPESSELSRKFRRTEEG 1962
QY 18 ----- 17
DB 1963 YYEAITAVELSRKDESEYELLRTKDKELLHWTKELEEKALAEGBKITPTFKPDK 2022
QY 18 ----- 17
DB 2023 IELSPSMEAPKIFERIQSVTQGSDAHFRVRVVGKPDPECEWKNGVKIERSDRIYVW 2082
QY 18 ----- 17
DB 2083 PEDNVCELVRDVTADSASINVKAINIAGETSSHAFLVQAKQLITFTQELQDVVAK 2142

QY 18 ----- 17
DB 2143 DTMAFECETSEPPVKVKKWYKDGMEVHEGDKYRMHSDRKHVFLSLTITDTSADAEYSCVL 2202
QY 18 ----- 17
DB 2203 VEDENVKTTAKLIVGAVVEFVKELQDIEVPESYGELECI VSPENIEGKWYHNDVELKS 2262
QY 18 ----- 17
DB 2263 NGKYTITRRGRQNLTVKDVTKEDQGEYSFVIDGKKTTCKLKMPPRIAILQLGSLDKVC 2322
QY 18 ----- 17
DB 2323 EGDIVOLEVKVLSVESVGVWMMKDGQEVOPSDRVHIVIDKQSHMLLIEDMTKEDAGNYSFT 2382
QY 18 ----- 17
DB 2383 IPALGLSTGRVSVYSDVITPLKDVNVIETGKAVLECKVSPDVTSVKWLNDQIKPD 2442
QY 18 ----- 17
DB 2443 DRVQAIKGTQKORLVINRTHASDEGPYKLIIVRVETNCLNLSVEKIKIIRGLDLTCTETQ 2502
QY 18 -----FMDK----- 21
DB 2503 NVVFEVELSHSGIDVNLNFKDKEIKPSSKYKIEAHGKIYKLTVLNMMKDDSEKTYFYAGE 2562
QY 22 ----- 21
DB 2563 NMTSGKLTVAGGAIKPLTDQTVAESQEA VPECEVANPDSKGEWLRDGLHPLTNIRSE 2622
QY 22 ----- 21
DB 2623 SDGKHRLIIIAATLDDIGEYTKVATSKTSAKLKVEAVKIKKLTNLTVTETQDQAVTV 2682
QY 22 ----- 21
DB 2683 ELTHPNVKGQWIKNGVVLSENEKYAISVKGTIYSLRIKNCALVDESYYGFRGLRGASA 2742
QY 22 ----- 21
DB 2743 RLHVETVKLIKKPKDVTALENATVAFESVSHDTPVKWFKHVSVEIKPSDKHRLVSEKV 2802
QY 22 ----- 21
DB 2803 HKMLQNLSPSDAGEYTA VQGLECKAKL FVETLHITKMNIEVPETKTASPECEVSHF 2862
QY 22 ----- 21
DB 2863 NVPSMWLKNNGVEIEMSEKFI VQGLHQLIIMNTSTEDSAEYTFVCGNDQVSATLTVTP 2922
QY 22 -----TH----- 24
DB 2923 IMITSMKLDINAEKDTITFEVTYVNEGISYKWLKNNGVEIKSTDKCQMRTKKLTHSLNR 2982
QY 25 ----- 24
DB 2983 NVHFGDAADYTFVAGKATSTATLYVEARHIEFRKHIDIKVLEKKRAMFECEVSEPDITV 3042
QY 25 ----- 24
DB 3043 QWMKDDQELQITDRIKIQEKYVHRLIIPSTRMSDAGKYTVVAGGNVSTAKLFEVGRDVR 3102
QY 25 ----- 24
DB 3103 IRSIKKEVQVIEKORAVVEFEVNEEDVDHAWYKDGIEINFQVQERHKYVVERRIHRMFIS 3162
QY 25 ----- 24
DB 3163 ETROSDAGEYTFVAGNRSSVTLYVNAPEPPVQLQELQVTVQSGKPARFCAMISGRPOP 3222

QY 25 ----- 24
Db 3223 KISWKEOQLSTGFKCKFLHDQOEVTLLEAFEDAAVYTCEAKNDYGVATTSASLSV 3282
QY 25 ----- 24
Db 3283 EPEVVSQDQEMPVYPPIITPLQDVTISGQPARQCRVSGDGLKVSWMYKDKKIKPSR 3342
QY 25 ----- 24
Db 3343 FFRMTQFEDTYQLEIAEAYPEDEGTTFVANNAGQVSSTANLSLEAPESILHERIEQEI 3402
QY 25 ----- 24
Db 3403 EMEWKEFSSFLSABEGLHSAELQLSKINETLELLSESPVYPTKFDSEKGTGPIFIKE 3462
QY 25 ----- 24
Db 3463 VSNADISMGDVATLSVTIGIPKPIQWFGVNGVLLTPSADYKFVFDGDHSLIILFTKLE 3522
QY 25 ----- 27
Db 3523 DEGEYTCMASNDYKGTICSAYLKINSKGEGHKDTETESAVAKSLEKLGPCPPHFLKELK 3582
QY 28 ----- 27
Db 3583 PIRCAOGLPAIFEYTVVGEPAPTVTFWFKENKQLCTSVIYITI IHNPNSGGTFIVNDPQRED 3642
QY 28 ----- 39
Db 3643 SGLYICKAENMLGESTCAABELLVLEDTMDTPCKRAKSTPEAPEDFPQTPLKGPAAVEAL 3702
QY 40 ----- 39
Db 3703 DSEQETATFVKDILKAALITEENQOLSYEHIAKANELSQLPLGAQELQSILOQDKLTP 3762
QY 40 ----- 41
Db 3763 ESTREFLCINGSIHFOPLKEPSNLQIIVQSOKTFSKEGILMPETQAVLSDEKIF 3822
QY 42 ----- 41
Db 3823 PSAMIEQINSLAVEPLKTLLAEPEGNYPOSSIEPPMHSYLTSVAEVLKKEKTSVSDTN 3882
QY 42 ----- 41
Db 3883 REORVTLQOEAOASALILSOSLAEGHVESLQSPDVMISQVNYEPLVPSEHSCHEGGKILI 3942
QY 42 ----- 43
Db 3943 ESANPLENAGDSAVRIEEGKSLRFPPLALEEKQVLLKEEHSNDVVMPPDQIIIESKREPVA 4002
QY 44 ----- 43
Db 4003 IKKQVEQGRDLLSKESLLSGIPPEORNLKIQICRALQAAVASEQPLFSEWLRNIEKV 4062
QY 44 ----- 43
Db 4063 EVEAVNITQEPRHIMCMWLYTSAKSVTEVTIIIEDVDPQMANLKMELRDALCAIYYEI 4122
QY 44 ----- 43
Db 4123 DILTAEGPRIQOAKTSLQEBMSFGSQKVEPITEPEVESKYLISTEEVSYNVQSRVK 4182
QY 44 ----- 43
Db 4183 YLDATPVTKGSAVVSDEKQDES LKPSKEESSESGETEATVKIQEAGGLIKEDG 4242
QY 44 ----- 43
Db 4243 PMIHTPLVDTVSEGDIVHLTTSITNAKEVNWYFENKLVPSDEKFKCLQDQNTYTLVIDK 4302
QY 44 ----- 43

Db 4303 VNTEDHQGEYVCEALNDSGKTATSAKLTVVKRAAPVIKRIEPLVALGHLAKFTCEIQS 4362
QY 44 ----- 43
Db 4363 APNVRFQWFKAGRIEYEDSKSTRSSKYISSLEILRTQVVVDGGEYCKASNEYGSVCTA 4422
QY 44 ----- 52
Db 4423 TLTVTVPGEKKVKRLLPERKPEKPEVVLKSVLRKRPEEEEPKVEPKLEKVKKPAVPE 4482
QY 53 ----- 52
Db 4483 PPPPKPVEEVPTVTKRKRKIPEPTKVPEIKPAIPLPAPEPKPKPEAEVKTIKPPVPEP 4542
QY 53 ----- 56
Db 4543 EPTPIAAPVTVPVVGKAAKAEAKPKGPIKGVPKPTSPPIEAERKRLRPGSGEK 4602
QY 57 ----- 56
Db 4603 PPDEAPFTYOLKAVPLKFVKEIKDIILTESEFVGSSAIFECLYSPSTAITTNMKDGSNIR 4662
QY 57 ----- 61
Db 4663 ESPKHRFIADGDKRKLHIIDVQLSDAGEYTCVLRNLGNKEKTSTAKLVEELPVRVKTL 4722
QY 62 ----- 61
Db 4723 EEVTVVGQPLYLSCELNKERDVMVRKDGKIVVEKFGRIVPGVIGLMRALTIINDADTDA 4782
QY 62 ----- 61
Db 4783 GTVTVTVENANNLECSKVVEVIRDWLVKPIRDQHVKPGTAIFACDIKADTPNIKWF 4842
QY 62 ----- 61
Db 4843 KGDEIPAEPNDKTEILROGNHLYLKIKNAMPEDIAEYAEIEGKRYPAKLTJGEREVEL 4902
QY 62 ----- 61
Db 4903 LKPIEDVTIYKEKESAFDAISEADIPGQWKLKGLLRPSPTCEIKAEKGKRFLLHKVK 4962
QY 62 ----- 61
Db 4963 LDQAGEVLYQALNAITAITLTVKEIELDFAVPLKDVTPERRQARPCVLTREANVWSK 5022
QY 62 ----- 61
Db 5023 GPDIIKSSDKFDIIADGKKHILVINDSQDFDEGYTAEVEGKKT SARLFTVGTIRLKF MSP 5082
QY 62 ----- 68
Db 5083 LEDQTVKEGTATFVGCLSHEKMHVWFKNDAKLHTSRVTLISSEGKTHKLEMKEVTLDD 5142
QY 69 ----- 68
Db 5143 ISOIKAOVKELSSAQLKVLADPYFTVKLHDKTAVEKDEITLKCEVSKDVPVKWPKDGE 5202
QY 69 ----- 68
Db 5203 EIVPSPKYSIKADGLRRLIKKADLKDKGEYVDCDGTDKTKANVTVEARLIEVEKPLYG 5262
QY 69 ----- 68
Db 5263 VEVFVGETAHEIELSEPDVHGQWKLKQPLTASPDCEIIEDGKKHILILHNCQLGMTGE 5322
QY 69 ----- 68
Db 5323 VSTQAAANAKSAA NLKV K ELP L I F I T P L S D V K V F E K D E A K F E C E V S R E P K T F R M L K G T Q E I 5382
QY 69 ----- 68

Db 5383 TGDRELIKDGTHKSMVKSAAFEDEAKYMEAEADKHTSGKLIIEGIRLKFLTPLKDYT 5442
QY 69 ----- 68
Db 5443 AKEKESAVTVLSDHNIIRVKFNQDRLHTTRSVSMQDEGKTHSITFKDLSIDTQSIR 5502
QY 69 -----DP----- 70
Db 5503 VEAMGMSSEAKLTVLEGGDPYFTGKLQDYGVEKDEVILQCEISKADAPVKWFKDGKEIKP 5562
QY 71 ----- 70
Db 5563 SKNAVIKTDGKKRMLILKALKASDIOGYTCDGTDKTSGLKDIEDREIKLVRPLHSVEYM 5622
QY 71 ----- 70
Db 5623 ETETARFETEISEDDIHANWKLGKALLQTPDCEIKEBGIHSLVLHNCRLDQTGGVDPQ 5682
QY 71 -----DP----- 70
Db 5683 AANVRSSAHLRVKPRVIGLLRPLKDVTVTAGETATFDCELSYEDIPVEWYLGKKLEPSD 5742
QY 71 ----- 70
Db 5743 KVVPRSEGVHTLRLRVKLEDAEYSCVETARTSGMLTVLEBEAVFTKLANIEVSETDT 5802
QY 71 ----- 70
Db 5803 ATAVLECEVSRENKVKWFKNGTEILKSKYIEIVADGRVKLVIHDCPTEDIKTYTCDAK 5862
QY 71 ----- 70
Db 5863 DFKTSCNLNVPPHVEFLRPLTDLOVREKEMARFECELSRENKVKWFKDGAIEIKKGRKY 5922
QY 71 ----- 70
Db 5923 DIISKAVRILVINKCLLDDEAEYSCEVETARTSGMLTVLEBEAVFTKLANIEVSETDT 5982
QY 71 ----- 70
Db 5983 IKLVCEVSKPGAIEVIWYKGDDEI IETGRVEIILTEGRKRLIVIONAHLEDAGNYNCRLPSS 6042
QY 71 ----- 70
Db 6043 RTDGKVKVHELAAEFISKPNLEILEGEKAEPVCSISKESFPVQWKRDDKTLESGBKYDV 6102
QY 71 -----EYKF 74
Db 6103 IADGKKRVLVVKDNLQDMGTVMVVGAAARAAHLTVIEKLRIVVPLKDRVKEQOEYVF 6162
QY 75 N----- 75
Db 6163 NCEVNTGAKAKWFRNEEAIFDSSKY IILQDVLVTLRLRDAHDDQANYVSLTNHRGE 6222
QY 76 ----- 75
Db 6223 NVKSAANLIVEEDRLIVEPLKDIETMEKKSVTFWCKVNRNLNVTLKWTKNGEVFPDNKV 6282
QY 76 ----- 75
Db 6283 SYRVDKYKHLMTIKDCGFPDEGEYIVTAGQDKSVAELLIIIEAPTEFVEHLEDQTVTEFDD 6342
QY 76 ----- 75
Db 6343 AVFSCQLSREKANVWYRNGRIEKGKKYKFEKDGSIHRLIIKDCRLDDECEYACGVEDR 6402
QY 76 ----- 75
Db 6403 KSRARLFVEEIPVEIIRPPQDILEAPGADVVFVLAELNKNKDEVOQLRNNMVVVGDKHQM 6462
QY 76 ----- 75
Db 6463 MSEGKIHLQICDIKPRQGEYRFTAKDKEARAKLELAAAPKIKTADQDLVVDVGKPLTM 6522

QY 76 ----- 75
Db 6523 VVPYDAYPKAAEWKFKENEPLSTKTIDTAEQTSFRILEAKKGDKGRKIVLONKHGKAE 6582
QY 76 ----- 75
Db 6583 GFINKLVIDVPGVPRNLEVTETFDGEVSLAWBEPLTDGSKIIGYVVERRDIKRTWVLA 6642
QY 76 ----- 75
Db 6643 TDRASCEFTVTGLQKGGVEYLFVRSARNRVGTGPVEDNPVEARSKYDVPGLPNTVI 6702
QY 76 ----- 75
Db 6703 TDVNRFGVSLTWEPPEYDGGAEITNYVIELRDKTSIRNDTAMTVRAEDLSATVTDVVEGO 6762
QY 76 ----- 75
Db 6763 EYSFRVRAQNRIGVGKPSAATPFVKVADPIERPSPVNLTSDDQTSQSSVQLKWEPLKDG 6822
QY 76 ----- 75
Db 6823 GSPILGYIIERCEEGKDNWIRCNMKLVPELYKYVTGLEKGNKYLRYVSAENKAGVSDPSE 6882
QY 76 ----- 75
Db 6883 ILGPLTADDAFVEPTMDLSAFKDGLEIVPNPITILVPSTGYPRPTATWCFGDKVLETDG 6942
QY 76 ----- 75
Db 6943 RVKMKTL SAYAELVISPERSDKGIYTLKLENRVKTIISGEIDVNVVIARPSAPKELKFGDI 7002
QY 76 ----- 75
Db 7003 TKDSVHLTWEPDDGGSPLTGYVVEKREVSRTKTKVMDVFVTDLEFVDPDLVQGEKYLEF 7062
QY 76 ----- 75
Db 7063 KVCARNKCGPEAYVDEPVNKMSTPATVPDPENVKWRDRTANSIFLTWDPKPDGGSRI 7122
QY 76 ----- 75
Db 7123 KGYIVERCPGRGSKWVACGEPVAETKMEVTGLEEGKWAYRVKTLNROGASKPSRPTBEI 7182
QY 76 ----- 75
Db 7183 QAVDTQEAPEIFLDVKLLAGLTVKAGTKIELPATVTGKPEPKITWTKADMILKQDKRITI 7242
QY 76 ----- 75
Db 7243 ENVPKSTVTIVDSKRSDTGYIIEAVNVCGRATAVVEVNVLDKPGPAAFDITDVTNES 7302
QY 76 ----- 75
Db 7303 CLLTWNPPRDDGGSKITNYVVERRATDSEVWHKLSSTVKOTNFKATKLIPNKEYIFRVA 7362
QY 76 ----- 75
Db 7363 ENMYGAGEPVQASPTAKYQFDPGPPTRLEPSDITKDAVTLTWCEPDGGSPTGYWV 7422
QY 76 ----- 75
Db 7423 ERLDPDTDKWYRCNKMPVKDITYRVKGLTNKKYRFRVLAENLAGPGKPSKSTEBILIKD 7482
QY 76 ----- 75
Db 7483 PIDPPWPBGKPTVKDVGKTSVRLNWTKEPHDGGAKIESYVIELMLKTGTDEWVRVAEGVPT 7542
QY 76 ----- 75
Db 7543 TQHLPLGLMEQGEYSFRVRAVNKAGESEPSFSDVLCREKLYPPSPRRLVINITKNT 7602

QY	76	-----	75
Db	7603	ADLKWTVPKDGSPITNIVYKRDYRKGQWQTVDTVKDKTCTVTLPLEGSLYVFRVAA	7662
QY	76	-----	75
Db	7663	ENAIQSDYTEIEDSVLAKDTFTTGPYPYALAVVDVTKRHVDLKBPPKNDGGRPIQRYV	7722
QY	76	-----	75
Db	7723	IEKKERLGRWVRKAGTAGDCNFRVTDVJEGTEVQFQVRAENEAGVGHPSPTLSIE	7782
QY	76	-----	75
Db	7783	DPTSPSPPLDLHVTDAGRKHIAIAWKPKPKNGGSPILGYHVEMCPVGTEKWRVNSRPI	7842
QY	76	-----	75
Db	7843	KDLKFRVEGVDPKEYVLRVRAVNAIGVSEPSEISENVVAKDPDCKPTIDLETHDIIVI	7902
QY	76	-----WYVDG-----	80
Db	7903	EGEKLSIPVFPRAVPVPTVSMHKDGKEVKASDRLTAKNDHISAHLEVPKSVRADAGIYTI	7962
QY	81	-----	80
Db	7963	TLENKLSATASINVKVIGLPGPKDIKASDITKSSCKLTWEPPEFDGGTPILHYLERR	8022
QY	81	-----	80
Db	8023	EAGRRTYIPVMGENKLSWTVDLIPNGEYFFRVKAVKVGGEYIELKNPVIAQDPKQP	8082
QY	81	-----VEVNA-----	86
Db	8083	PDPVVDVEVHNPTAEAMTITWKPPLYDGGSKIMGYIIIEKIAKEERWKRCNEHLVPILTY	8142
QY	87	-----	86
Db	8143	TAKGLEEGKEYQFVRRAENAGISEPSRATPPTKAVDPIDAPKVILRTSLVARGDEIAL	8202
QY	87	-----	86
Db	8203	DASISGSPYTIITWIKDENVIVPEEIKKRAAPLVRRRKGEVQEEPPVLTQRLSIDNS	8262
QY	87	-----	86
Db	8263	KGESQLVRDSLRPDHGLYMIKVENHDGIAKAPCTVSVLDTPGPPINFVFEDIRKTSVL	8322
QY	87	-----KTKP-----	90
Db	8323	CKWEPPLDGGSEIINVTLEKDKTKPDSEWIVVTSTLRHCKYSVTKLIBGKEYLFRVRA	8382
QY	91	-----	90
Db	8383	ENRFGPPPCVSKPLVAKDPGPPDAPDKPIVEDVTSNSMLVKWNEPKDNGSPILGYWLE	8442
QY	91	-----	90
Db	8443	KREVNSTHNSRVNKSLLNALKANVDGLLEGLTVFVRVCAENAGPGKFSPPDPKTAHDP	8502
QY	91	-----	90
Db	8503	ISPPGPPIPRVTDTSSTTIELEWEPPAFNGGGEIVGYFDKQLVGTNKNWRCCTEKMIKVR	8562
QY	91	-----	90
Db	8563	QYTVKEIREGADYKLRVSAVNAAGEGPPGETQPVTVAEPOEPFAVELDVSVKGGIQUIMAG	8622
QY	91	-----	90
Db	8623	KTLRIPAVVTGRPVTKVWTKEGELDKDRVIDNVGTKSELIIKDALRKDHGRYVITAT	8682
QY	91	-----	90

Db	8683	NSCGSKFAARVEVFDVPGPVLDLKPVVNTNRKMCLLNWSDEDDGGSEITGFIIERKDAK	8742
QY	91	-----	90
Db	8743	MHTWROPDIETERSKCDITGLLEGOEYKFRVIANKKFGCGPPVEIGPILAVDPLGPPTSPE	8802
QY	91	-----	90
Db	8803	RLTYTERQSTITLDWKEPRSGSGPIQGYIIIEKRHRDKPDFERVKRKLCTPTTSFLVENL	8862
QY	91	-----	90
Db	8863	DEHQMEFRVKAYNEIGESEPSPLPLNVVIODDEVPPTIKRLSVRGDTIKVKAGEPVHIP	8922
QY	91	-----	90
Db	8923	ADVTGLPMPKIEWSKNETVIEKPTDALQITKEEVSRESEKTELSPKAVREDKGTYYTVA	8982
QY	91	-----	90
Db	8983	SNRLGSFVRNVHVEYDRPSPPRNLAVIDIKAESCYLTWDAPLDNGGSEITHYVIDKRDA	9042
QY	91	-----	90
Db	9043	SRKKAWEETNTAVEKRYGIWKLIPNGQYEFVRVAVNKYGISDECKSKVVIQDPYRLP	9102
QY	91	-----	90
Db	9103	GPPGPKVLARTKGLSVLWSWTPPLDNGSPITGYWLEKREEGSPYWSRVSRAPITKVGLK	9162
QY	91	-----	90
Db	9163	GVEFVNRLLLEGVYOFRAMAINAAGIGPPSESDPEVAGDPIFPPOGPPSCPEVKDKTKS	9222
QY	91	-----	90
Db	9223	SISLGHKPPAKDGGSPILGYIVEMQEBGTTDWKRVNEFDKLITTCVCPVNLKELRYRF	9282
QY	91	-----	90
Db	9283	RVKAVNAGESEPSDTTGEIPATDIOEEPEVFIDIGAQCVCVKAGSQRIPAVIKGRPT	9342
QY	91	-----	90
Db	9343	PKSSEFPDGAKKAMKGVHDIPEDAOLEAENS SVIIIECKRSHTGKYSITAKNKAGQ	9402
QY	91	-----	90
Db	9403	KTANCRVKVMDVPGPKDLAVSDITRGSRLSKWMPDDGGDRIKGVVIEKRTIDGKAWT	9462
QY	91	-----REEQY-----	95
Db	9463	KVNPDCGSTTVVVPDLLSEQOYFFRVRAENRFGIGPPVETIQRTARDPIYPPDPIKLK	9522
QY	96	-----	95
Db	9523	IGLITKNTVHLSWPKXNDGSGSPVTHYIVECLAWDPTGTTKEAWRQCNKRDRVELOFTVE	9582
QY	96	-----	95
Db	9583	DLVEGGEYEFVRKAVNAAGVSKPSATVGPCQCORPMDPPSIDLKEFMEVEEGTNVIVAK	9642
QY	96	-----	95
Db	9643	IKGVPPFTLTFWKAPPKPDNKEPVLVDTHVNKLVDVDTCTLVIPQSRSDTGLYTITAV	9702
QY	96	-----	95
Db	9703	NNLGASKEMRLNVLRPGPPVGPPIKFSVSADQMTLSWFPDDGGSKITNVIKREA	9762
QY	96	-----	95

Db 9763 NRKTVHVSSEPKECTYITPKLLEGHEVYFRIMAQNKYIGIBLSEPETARNLFSVPGA 9822
QY 96 ----- 95
Db 9823 PDKPTVSSYTRNSMTVNWEEPEYDGGSPYGYWLEKMDTTSKRWRVNRDPIKAMTLGVS 9882
QY 96 ----- 95
Db 9883 YKVTGLIEGSDYQFRVYAINAAGVPASLPSPATARDPIAPPGPPPKVTDWTKSSADL 9942
QY 96 ----- 95
Db 9943 EWSPPPLKOGSKVTGVIYVEYKEEGKEWEKDKKEVGRGKLVVTGLKEGAFYKFRVSAYN 10002
QY 96 ----- 95
Db 10003 IAGIGEPEVTDVIEKDRVSPDLQDASVRDRIVVHAGGVIIRIAYVSGKPPFTVTWN 10062
QY 96 ----- 95
Db 10063 MNERTLPQEAETIETTAISSSMVINKQCRSHQGVYSLAKNEAGERKKTIIVDVLDPGPV 10122
QY 96 ----- 95
Db 10123 GTPFLAHLNLTNSCKLTWFSPEDDGSPITNYVIEKRESRRRAWTPVTYVTRQNATVQG 10182
QY 96 ----- 95
Db 10183 LIOGRKAYFPRIAAENSIGMPFVETSEALVIREPITVPERPELEVKEVTKNTVTLTWNP 10242
QY 96 ----- 95
Db 10243 PKYDGSSEIYNVLESRLIGTEKFHKVTNDLLSRKTYVKLKEGDTVEYRVSAVNIQO 10302
QY 96 ----- 95
Db 10303 GKPSFCTKPTCKDELAPTLHLDFRDLTIRVGEAFALTGRYSGKPKVSWFDEADY 10362
QY 96 ----- NST----- 98
Db 10363 LEDDRTHIKTATLAEKIKAKRSDSGKYCVVENSIGSRKGFQCVNVVDHPPGPV 10422
QY 99 ----- 98
Db 10423 SFDEVTKDYVSWKPPLDGSGKITNYLIEKKEVGDVWMPVTSASAKTTCKVSKLLEG 10482
QY 99 ----- 98
Db 10483 KDYIFRIHAENLYGISDPLVSDSMKAKDRFRVPDAPQPIVTEVTKDSALVTWNNKPHDGG 10542
QY 99 ----- 98
Db 10543 KPITNYILEKRETMKRWARVTKDPIHPTKFRVPDLLEGCOYEFPRVSAEENIGIDPSP 10602
QY 99 ----- 98
Db 10603 PSKVPFAKDPIAKSPVPVPEAIDTTCNSVOLTWQPRHDGSKILGYIYVEYQKVGDSEW 10662
QY 99 ----- YRVSV----- 104
Db 10663 RRANHTPESCPTKYKVTGLRDGOTYKFRVLAVNAAGESDPAHVPEPVLVKDRLEPPELI 10722
QY 105 ----- 104
Db 10723 LDANMARQHIKVGDTLRLSAILKGPVPPKVTWKKEDRDAPTKARIDVTPVGSKLEIRNA 10782
QY 105 ----- LTV----- 107
Db 10783 AHEDGGIYSLTVENPAGSKTVSVKVLVDKPGPRDLEVSEIRKDCSYLTWKEPLDDGGS 10842
QY 108 ----- 107
Db 10843 VITNTVVERRDVASAQWSPLSATSKKSHFAKHLNEGNOYLFRVAAENYOYGRGPPVETPK 10902

QY 108 ----- LH----- QDWL-- 113
Db 10903 PIKALDPLHPPGPKDLHHVDVDKTEVSLVWNKPRDGGSPITGYLVYEQEGTQDWIKF 10962
QY 114 ----- 113
Db 10963 KTVTNLECVVTGLQOGKTYRFRVKAENIVGLGLPDTTPIECOEKLVPPSVBELVKLIBG 11022
QY 114 ----- 113
Db 11023 LUVKAGTTVRFPAILIRGVPVPTAKWTTDGSEIKTDEHYTVETDNFSSVLTIKNCLARDTG 11082
QY 114 ----- 113
Db 11083 EYQITVSNAAGSKTVAVHLTVLDVPCPTGPIINILDVTPHEMTISWQPKDDGGSPVINY 11142
QY 114 ----- 113
Db 11143 IVEKQDTRKDTWGVVSSGSSKTKLIPHLOKGEYVFRVRAENKICGVGPPLDSTPTVAKH 11202
QY 114 ----- 113
Db 11203 KFSPPSPCKPVVTDITENAATVSWTLPKSDGSGSPITGYMERREVTGKWRVNKPTIAD 11262
QY 114 ----- 113
Db 11263 LKFRVTGLYEGNTYBFRVFAENLAGLSKSPSSDPIKACRPKIPGPPINPKLKOKSRT 11322
QY 114 ----- 113
Db 11323 ADLVWTKPLSDGSGPILGYVVECOQPGTAQMNRIKDELIROCAFVPGLEGNEYRFR 11382
QY 114 ----- 113
Db 11383 KAANIVGESEPRELAESVIAKDILHPPEVELDVTCDVITVRVGOTIRILARVKGRPEPD 11442
QY 114 ----- 113
Db 11443 ITWTKEGVLVREKRVDLIQLDLPVELQIKEAVRADHGKYIISAKNSSCHAOGSAIVNVL 11502
QY 114 ----- 113
Db 11503 DRPGCQNLKVTNVTKENTISWENPLDNGGSEITNFIVEYRKNQKGSIVASDVTKRL 11562
QY 114 ----- 113
Db 11563 IKANLLANNEYFRVCAENKVCVGPTIETKTPILAINPIDRGPENLHIAKGTFFVL 11622
QY 114 ----- 113
Db 11623 KWRRPDYDGGSPNLSYHVERRLKGSDDHVHUKGSIKETHYVNDRCVENQIYEFVRVQTKN 11682
QY 114 ----- 113
Db 11683 EGGSDWVKTEBVVVKEDLQKPVLDLKLGLSVLTAKGDTIRLEAGVRGKPPPEVAVTKDK 11742
QY 114 ----- 113
Db 11743 DATDLTRSPRKIDTRADSSKFSLTKAKRSDGKYVVVTTATNTAGSFVAYATVNVLDKPGP 11802
QY 114 ----- 113
Db 11803 VRNLKIVDVSSDRCTVCWDPPDDGGCEIQNYILEKCETKRWVWSTYSATVLTPGTTVTR 11862
QY 114 ----- 113
Db 11863 LIEGNEYIFRVAENKIGTPPTESKPIAKTKYDKGRDPDPPEVTKVSKSEMTVWVWNP 11922
QY 114 ----- 113
Db 11923 EYDGGKSTGYFLEKKEKHSTRVWPVNKSAIPERRMKVONLLUPDHEYQFRVKAENEIGIG 11982

QY 114 ----- 113
Db 11983 EPSLSPRVVAKDIEPPGPPTNFRVVDTKHSITLWGKXPYDGGAPIIGYVVMRPKI 12042
QY 114 ----- 113
Db 12043 ADASPDGKRCNAALVRKEFTVTSLDENQEYEFVRCANQVIGRPAELKEAIKPK 12102
QY 114 ----- 113
Db 12103 ILEPPEIDLASMRKLIVIRAGCPIRLFAIVRGPAPKVTWRKVGDVNVVRKQGVLDVT 12162
QY 114 ----- 113
Db 12163 MAFIVIPNSTRDSDGKYSLTLVNPAGEKAVFVNVRLDTPGPVSDLKVSDVTKTSCHVSW 12222
QY 114 ----- 113
Db 12223 APPENDGGSQVTHYIVEKREARKTWSTVTPVKKTSFHVTLNLPNGEYFRVTAVNEYG 12282
QY 114 ----- 113
Db 12283 PGVPTDVKPVLASDPLSEDPKPRKLEATEMKNATLAWLPLRGGAKIDGYIISYRE 12342
QY 114 ----- 113
Db 12343 EEQPADRWTEYSVVKDLSLVVTGLKEGKYYKFRVAARNVGVSLPREAGYEAKEQLLP 12402
QY 114 ----- 113
Db 12403 PKILMPQITIKAGKLRIEAHVYGPHTCKWKKGDEVVTSVSHLAVHKADSSILIK 12462
QY 114 ----- 113
Db 12463 DVTRKDSGYSLAENSSGTDQKIKVVMDAPGPPQPPEDISDIDADACSLSWHIPLD 12522
QY 114 ----- 113
Db 12523 GGSNITNYIVEKCDVSRGDNVTALASVTKTSCRVGKLIIPGQEIFRVAENRFGISEPLT 12582
QY 114 ----- 113
Db 12583 SPKMAQFPFGVPSEPKNARVTKVNDKICFVAMDRPDSGGSPIIGYLIBERKERNLLW 12642
QY 114 ----- 113
Db 12643 KANDTLVRSTPCAGLVEGLEYSFRIYALNKAGSSPPSKPTXYVTARMPVDPGKPEVI 12702
QY 114 ----- 113
Db 12703 DVTKSTVSLIWARPKHDGSKIIGYFVEACKLPDGKWRNCNTAPHOIPQEEYTATGLEEK 12762
QY 114 ----- 113
Db 12763 AQYQFRAIARTAVNISPPSEDPVTIILAENVPRIDLVSAMKSLTAVKAGTNVCLDATV 12822
QY 114 ----- 113
Db 12823 FGKPMPTVSWKDGTLKPAEGIKMAMQORNLCITLFLSVNRKSDGYTTIATENSSGSKSA 12882
QY 114 ----- 113
Db 12883 TIKLVLDKCPPASVKINKMYSDRAMLSWEPPLDGGSEITNYIVDKRETSRPNWAQVS 12942
QY 114 ----- 113
Db 12943 ATVPITSCSVEKTEGHEYQFICAENKYGVDVFTPEALAKNPYDPPGRCDDPVISNI 13002
QY 114 ----- 113
Db 13003 TKDHMTVSWKPPADGGSPITGYLLEKRETAQVNWTKVNRKPIIERTLKATGLOEGTEYE 13062
QY 114 ----- 113

Db 13063 FRYTAINKAGPKPSDASKAAYARDPOYPPAPAPFVKVYDTRSSVSLSMGKPAYDGGSP 13122
QY 114 ----- 113
Db 13123 IIGYLVEVKRADSDNVVRCNLQNLOKTRFEVTGLMEDTOYQFRYAVANKIGYSDPSDVP 13182
QY 114 ----- 113
Db 13183 DKHYPKDILIPPEGEHDADLRKTLILRAGVTMRLYVPVKRPPPKITWSPKNWLRDRIG 13242
QY 114 ----- 113
Db 13243 LDIKSTDFDFLRCENVNKYDAGKYILTLENSCGKKEYTIWVKVLDTPGPPINVTVKEIS 13302
QY 114 ----- 113
Db 13303 KDSAYVTWEPPIIDGSPPIINVVQKRDAERKSMSTVTTECKSTFRVPNLEGKSYFFR 13362
QY 114 ----- 113
Db 13363 VFAENEYIGDGPETRDVAVKASQTPGPVVDLKVRSVSKSSSIGWKPHDGGSRILGYV 13422
QY 114 ----- 118
Db 13423 VDFLTEENKQWVMKSLSLQYSAKDLTEGKEYTFRVSAENENGEGTPTSEITVWARDVVA 13482
QY 119 ----- 118
Db 13483 PDLKGLPDLCLAKENSFRLKPIKGPAPSVSMKKGEDPLATDTRVSESSAVNTT 13542
QY 119 ----- 118
Db 13543 LIVVDCOKSAGKVTITLKNVAGTKEGTISIKVVGKPGIPTGPIKFDEVTAEAMTLKWP 13602
QY 119 ----- 118
Db 13603 PKDDGGSEITNYILEKRDVNNKWNVCASAVOKTTFRVTRLHEGMEYTFRVSAENKYGVG 13662
QY 119 ----- 118
Db 13663 EGLSEPIVARHPFDPVDPAPPNNIVDVHRHDSVSLTWTDPKKTGGSPITGYHLEFKERN 13722
QY 119 ----- 118
Db 13723 LLKWRANKTPIRMDFKVTGLTEGLEYEFRVMAINLACVGKPSLPSEPVVALDPIIDPGK 13782
QY 119 ----- 118
Db 13783 PEVINITRNSVTLIWTEPKYDGGHKLTYIVEKRDLPFSKMMKANHVNPCEAFTVTDLV 13842
QY 119 ----- 118
Db 13843 EGGKYEFRIAKNTAGAISAPSESTETIICKDEYEAPTIVLDPTIKDGLTIKAGDTIVLN 13902
QY 119 ----- 118
Db 13903 AISILGKPLPKSSWSKAGKDIRPSDITQITSTPTSSMLTIKYATRKDAGEYTIATNPF 13962
QY 119 ----- 118
Db 13963 TKVEHVKTVDVPGPPGPEISNVSAEKATLTWTPPLEDGGSPIKSYILEKRETSRLW 14022
QY 119 ----- 118
Db 14023 TVWSEDIQSCRHVATKLIQGNIEYIFRVSAVNHYGEPVQSEPVKMWDRFGPPGPEKPE 14082
QY 119 ----- 124
Db 14083 VSNYTKNTATVSWKRPVDDGGSEITGYHVERREKSLRWRAITPVSDDLCKVTLGQEG 14142
QY 125 ----- 124

Db 14143 STYEFVSAENRAGIGPPSEASDSVLMKDAAYPPGPPSPHVTDTTKKSASLAWGKPHYD 14202
QY 125 ----- 124
Db 14203 GGLEITGVVVEHQVGDEAWIKDTGTGALRITQFVVVDPDLQTKENYFRISAINDAGVGE 14262
QY 125 ----- KALPAP----- 130
Db 14263 AVIPDVEIVEREMAPDFELDAELRRTLIVVRAGLSIRIFVPIKGRPAPEVTWTNDINLKN 14322
QY 131 ----- 130
Db 14323 RANIENTESFTLLIIPECNRYDTGKFMVTIENPAGKSGFVNVRLDTPGPVLNLRPTDI 14382
QY 131 ----- IEK----- 133
Db 14383 TKDSVTLHNDLPLJDGSRITNIVEKREATRKSYSATTKCHKTYKVTGLSEGCEYFF 14442
QY 134 ----- 133
Db 14443 RVMAENEYIGEPTETEPVKASEAPSPDLSNIMDITKSTVSLAWPKPKHDGSGKITGY 14502
QY 134 ----- KALPAP----- 133
Db 14503 VIEAQRKSGDQWTHITTVKGLECVVRNLTEGEYTFQVMVNSAGRSAPRESRPVIVKEQ 14562
QY 134 ----- 133
Db 14563 TMLPELDLRGIYOKLVIKAGDNKVEIPVLGRPKPTVWKKGQDILKQTORVNFETTAT 14622
QY 134 ----- 133
Db 14623 STILNINECVRSDBGPYPLTARNIVEGVDVITIQVHDIPGPTGPIKFEVSSDEVTS 14682
QY 134 ----- 133
Db 14683 WDPENDGGVPISNYVVEMRQDSTTWVELATTVIRTTYKATRLTTGLEYPQFVKAQNR 14742
QY 134 ----- 133
Db 14743 GVGPGITSAWIVANYPFKVPGPGTQVAVTKDSMTISWHEPLSDGSGPILGYHYVERKE 14802
QY 134 ----- TISKA----- 138
Db 14803 RRGILMQTVSKALVPGNIFKSSGLTDGIAEFVRIAENMAGSKSPSEPMALDPIDP 14862
QY 139 ----- 138
Db 14863 PGKPVPLNTRHTVTLKWAKEPYTGFKITSYIVEKRDLPNGRWLKANFSNILENEFTVS 14922
QY 139 ----- 138
Db 14923 GLTEDAAEFVRIAENMAGSKSPSEPMALDPIDP 14982
QY 139 ----- 138
Db 14983 RLEADVSRPPTMEWSKDGKEGTAKLEIKIADFSTNLVKNKSTRDSGAYTLTATNP 15042
QY 139 ----- 138
Db 15043 GGFAKHIFNVKVLDRPGPPGPLAYTEVISEKCVLSWFPPDDGGAKIDHYIVQKRETSR 15102
QY 139 ----- 138
Db 15103 LAWINVASEVQVTKLVTKLLKGNEYIFRMAVKNYGVGEPLSEBVLAVNPGPPDPK 15162
QY 139 ----- 138
Db 15163 NPEVTITKDSMVWCVGHPDSDGGSEIINYIIVERDRKAGQWIKCNKKTLDRLYKVSGL 15222
QY 139 ----- 138
Db 15223 TEGHEVEFRIMAENAGISAPSTSPFFYKACDTVFKPGPPGNPRVLDTSRISISIAWNKP 15282

QY 139 ----- 138
Db 15283 IYDGSSEITGYMVEIALPEDEMQIVTPPAGLAKATSYITGLTENQEKIRIYAMNSEGL 15342
QY 139 ----- 138
Db 15343 GEPALVPGTPKAEDRMLPEIELDADLRKVVTIRACCTLRFPVPIKGRPDPEVKWARDHG 15402
QY 139 ----- 138
Db 15403 ESLDKASIESASSYTLIIVGNVNRFDGKYILTVENSNGSKSAFVNVRLDTPCPQDLK 15462
QY 139 ----- 138
Db 15463 VKEVTKTSVTLTWDPPLLDGGSKIKNYIVEKRESTRKAYSTVATNCHKTSWKVDQLQEGC 15522
QY 139 ----- 138
Db 15523 SYFVRVLAENEYIGLPAETAESVKASERPLPPGKITLMDVTRNSVLSWKEPEHDGSR 15582
QY 139 ----- 138
Db 15583 ILGYIVEMQTKGSKWATCATVKVTEATITGLIQGEEYSFRVSAQNEKIGISDPRQLSPV 15642
QY 139 ----- 138
Db 15643 IAKDLVIPPAFKLLFNTFTVLAGEDLKVDPVFFIGRPTPAVTHWIKDNVPLKQTRVNAEST 15702
QY 139 ----- 138
Db 15703 ENNSLLTIKDACREDVGHVYVWKLNSAGEALETNLVILDKPGPTGPKVMDEVTADSIT 15762
QY 139 ----- 138
Db 15763 LSWGPPKYDGGSSINNYIVEKRDSTTTWQIVSATVARTTIKACRLTKCEYOFRIAAEN 15822
QY 139 ----- 138
Db 15823 RYCKSTYLANSEPTVAQYPFKVPGPGTVPVTLSSRDSMEVQWNEPISDGGSRVIGYHLER 15882
QY 139 ----- 138
Db 15883 KERNSILWVKNLKTPIQTKFKTTGLEGEVEYEFVSAENIVGIGKPSKVCYVARDPC 15942
QY 139 ----- 138
Db 15943 DPPGRPEAIIVTRNSVTLOWKKPTVDGSGKITGYIVEKKELPEGRWKKASFTNIIDTHFE 16002
QY 139 ----- 138
Db 16003 VTGLVEDHRYEFVRIAENAGVSEFSESTGALTARDEVDPRI SMDPKYKDTIVVHAGE 16062
QY 139 ----- 138
Db 16063 SPKVDADIYKPIPTIQWIKGQELSNARLEIKSTDFATSLSKDAVRVDSGNYILKAK 16122
QY 139 ----- 138
Db 16123 NVAGERSVTNVKVLDRPGPPGPGPVVISCVTAEKCTLAWKPLQDGGSDIINIIVERRET 16182
QY 139 ----- 138
Db 16183 SRLVWTVADVQTLCKVTKLLEGNEYTFRIMAVKNYGVGEPLSEBVLAVNPGPPDPK 16242
QY 139 ----- 138
Db 16243 PKAPEVTTVTKDSMIWVWERPASDGGSEITLGYVLEKROKEGIRWTRCHKRLIGELRLRVT 16302
QY 139 ----- 138
Db 16303 GLIENHDIYEFVRIAENAGLSEPPSPSAQKACDPIYKPPNNPKVIDITRSSVFLSWS 16362

QY 139 ----- 138
Db 16363 KPIYDGGCEIOGYIVERKDVNVGEWTCTPTTGINKNTNIEVEKLEKHEYNFRICAINKA 16422
QY 139 ----- 138
Db 16423 GVGEHADVPGPPIIVEEKLEAPDIDLLELRKIIINIRAGGSURLFVPVINGRPTPEVKWGV 16482
QY 139 ----- 138
Db 16483 DGEIRDAALIDVTSFTSLVDLNVNRDYSKYITLTLENSGTSKSAFVTVRVLDTPSPVN 16542
QY 139 ----- 138
Db 16543 LKVTEITKDSYSITWEPELLDGGSKIRNYIVEKREATRKSYAAVVTNCHKNKSKIDQLQE 16602
QY 139 ----- 138
Db 16603 GCSYFRVTAENEYIGLPAQTADPIKVAEVPQPPGKITVDVDRNSVLSWTKPEHDGG 16662
QY 139 ----- 138
Db 16663 SKIIQYIVEMQAKHSEKSEKARVKSLOAVITNLTQGEYLFPRVAVNNEKGRSDPRSLAV 16722
QY 139 ----- 138
Db 16723 PIVAKDLVIEPDVKPAFSSYSVQVQDLMKEVPISGRPKPTITWTKDGLPLKQTRINVT 16782
QY 139 ----- 138
Db 16783 DSDLTLTISIKETHKDDGGGYGITVANVVGOKTASIEIVTLDKDPKPGVKFDDVSAES 16842
QY 139 ----- 138
Db 16843 ITLSWNPPLYTGGCOITNYIVQKRDTTWTWDDVVSATVARTTLKVKLTKTEYQFRIFA 16902
QY 139 ----- 138
Db 16903 ENRYQSFALSDPIVAQPYKPGPGTPEFATAISKDSMVQIOWHEPVNNGSPVIGVHL 16962
QY 139 ----- 138
Db 16963 ERKERNILWTKVANKTIIHDTPQAKNLEEGIEYEFVRYAENIVGVGKASKNSECYVARD 17022
QY 139 ----- 138
Db 17023 PCDPPGPPEIMVXRNEITLQWTKPVYDGGSMITGYIVEKKRDLDPGRWMAKSFNTNIETQ 17082
QY 139 ----- 138
Db 17083 FTVSGLTEDQRYEFRVIKNAAGAIKSPSDSTGPITAKDEVELPRISMDPKFRDTIVVNA 17142
QY 139 ----- 138
Db 17143 GETFRLEADVHGKPLPTIEWLRGDKIEESARCEIKNTDFKALLIVKDAIRIDGGQYILR 17202
QY 139 ----- 138
Db 17203 ASNVAGSKSPFVNVKVLDRPGPPEGPVQVGTSEKCSLTWSPPLQDGGSDISHYVVEKR 17262
QY 139 ----- 138
Db 17263 ETSRLATVAVASEVVTNSLKVTKLLEGNEYVFRIMAVNKNYGVGELESAPVIMKNPFVLP 17322
QY 139 ----- 138
Db 17323 GPPKSLEVTNIAKDSMTVCWNRPSDGGSEIIGYIVEKRRDSGIWIKCNKRRTDLRLR 17382
QY 139 ----- 138
Db 17383 VTGLTEDEYEFVRSNAAGVGPSPATVYKACDPVKPGPPTNAHIVDTTKNSITLA 17442
QY 139 ----- 138

Db 17443 WGKPIYDGGSEILGVVVEICKADEEWOIVTPQTGLRVTRFEISKLTEHQEYKIRVCALN 17502
QY 139 ----- 138
Db 17503 KVGLEATSVPGTVKPEDKLEAPELDLSELKGIIVRAGGSARIHIPFKGRPMPEITWS 17562
QY 139 ----- 138
Db 17563 REEGFTDKVOIEKGVNTQLSIDCNRNDAGKYITLKLENSGGSKSAFVTVKVLDTPGPP 17622
QY 139 ----- 138
Db 17623 QNLAVKEVRKDSAFVLWEPPIIIDGGAVKVNYVIDKRESTRKAYANVSSKCSKTSFKVENL 17682
QY 139 ----- 138
Db 17683 TEGAIYYFRVMAENEFVGVPVETVDAYKAAEPSPPGKVTLTDVTSOTSASLMWEKPEHD 17742
QY 139 ----- 138
Db 17743 GGSRVLYVVEWQPKGTEKWSIVAESKVCNAVVTGLSSGQEQYQFRVKAYNEKGSDDPRVL 17802
QY 139 ----- 138
Db 17803 GVPVIKDLTIQPSLKLFPNTYSIQAGEDLKIEIPVIGRPRNISWVKDGEPLKQTRVN 17862
QY 139 ----- 138
Db 17863 VEETATSTVLHIKEGNKDDFGKTYVTATNSAGTATENLSVILEKPPVPGVFRFDEYSA 17922
QY 139 ----- 138
Db 17923 DFVVISHEPPAYTGGCOISNVIIVEKRDTTTTWHMVSATVARTTIKIKLTKTEYQFRI 17982
QY 139 ----- 138
Db 17983 FAENRYGKSAPLDSKAVIVQYPFKEPPGTPFTVISKQMLVQWHEPVDNDGGTKIIGY 18042
QY 139 ----- 138
Db 18043 HLEQEKNSILWKLNKTPIQDTKFTGLDLEGEYEFKVAENIVGIGKFSKSECFVA 18102
QY 139 ----- 138
Db 18103 RDCDPPGPRPEAIVITRNNVTLKWKKPAYDGGSKITGYIVEKKDLDPGRWMAKSFNTNLE 18162
QY 139 ----- 138
Db 18163 TEFTVSGLVEDQRYEFRVIARNAAGNFSEPSDSSGAIARDEIDAPNASLDPKYKDVIVV 18222
QY 139 ----- 138
Db 18223 HAGETVLEADIRGKPIPDVWVMSKDGKLEETAARMEIKSTQKTTLVVKDCIRTDGGQY 18282
QY 139 ----- 138
Db 18283 ILKLSNVTGTSIPITVTVKVLDRPSPGPEPLKVTGVTAEKCYLANWNPLODGGANISHXII 18342
QY 139 ----- 138
Db 18343 EKRETSRLSWTQVSTEVQALNXYKVTKLPGNEYIFRVMVAVNKNYGIGEGLESQPVTAACNPY 18402
QY 139 ----- 138
Db 18403 KPPGPPSTPEVSATIKDSMWVTWARPVDDGGTEIEGYILEKRDKEGVRWTKCNKKTLDL 18462
QY 139 ----- 148
Db 18463 RLVRTGLTEGHSEYFRVAENAAGVGESEPSVFRACDALYPPGPPSNPKVTDTSRSV 18522
QY 149 ----- 148

Db 18523 SLAWSKPIYDGGAPVKGYYVEVKEAAADEWTTCTPTGLOGKQFTVTKLKENTENFRIC 18582
Qy 149 -----TLP----- 151
 |||
Db 18583 AINSEGVCEPATLPGSVVAQERIEBPPELELDADLRKVVVLRSATLRLFTVTKRPEPEV 18642
Qy 152 ----- 151
Db 18643 KWKAEGILTDRAGIEVTSSTMLVIDNVTRFDSGRYNLTLENNSGSKTAFVNVRLDSP 18702
Qy 152 ----- 151
Db 18703 SAPVNLITREVKDSVTLSEWPEPLIDGAKITNYIVEKRETRKAYATIINCTKTTFRI 18762
Qy 152 ----- 151
Db 18763 ENLQEGCSYFRVLASNEYGIGLPAETEPVKYSEPPPLPGRVTLVDVTRNTATIKWEKP 18822
Qy 152 ----- 151
Db 18823 ESDGSKITGVVEMOTKGSEKSTCTQVKTLEATISGLTAGEEYVFRVAAVNEKGRSDP 18882
Qy 152 ----- 151
Db 18883 RQLGVPVIARDIEIKPSVELPFHTFNVKAREQLKIDVPFKGRPOATVNRKDGQTLKETT 18942
Qy 152 ----- 151
Db 18943 RVNVSSKTVTSLSIKEASKEDVGYELCVSNSAGSITVPITIVLDRPGPPGPIRIDEV 19002
Qy 152 ----- 151
Db 19003 SCDSITISWNPPEYDGGCOISNYIVEKKETTSTTHIVSOAVARTSIKIVRLATGSEYQF 19062
Qy 152 ----- 151
Db 19063 RVCAENRYGKSSYESSAVVAEYFPFPPGPGTPKVVHATKSTMLVTWQVPVNDGGSRI 19122
Qy 152 ----- 151
Db 19123 GYHLEYKERSILWSKANKILIADTVKVSGLDEGLMEYRYVAENIAGIGKSKCEPV 19182
Qy 152 PSRD----- 155
 |:|
Db 19183 PARDCDPPGQPEVNTITRSVSLKWSKPHYDGGAKITGYIVERRELDPGRWLKCNVTNI 19242
Qy 156 ----- 155
Db 19243 QETYFEVTELTEDQRYEFRVFARNAADSVSEPSSESTGPIIIVKODVEPPRVMMDVKPRDVI 19302
Qy 156 ----- 155
Db 19303 VKRAGEVLKINADIAGRLPLVISWAKDGIIEERARTEIISTDNHTLLTVKDCIRDTGQ 19362
Qy 156 ----- 155
Db 19363 YVLTLNAGTRSAVNCKVLDKPGPPAGPLEINGLTAECKSLSWGRPOEGGADIDYXH 19422
Qy 156 ----- 155
Db 19423 RKKRETSHLAWTICEGELQWTSCKVTLLKGNEYIFRVTGVNKYGVGEPLSVAIKALDP 19482
Qy 156 ----- 155
Db 19483 FTVPSPTSLEITSVTKESMTLCWSPESDGGSEISGYIIEERKNSLRWRVNRKPKPYVD 19542
Qy 156 ----- 155
Db 19543 LRVKSTGLREGCEYRYVAENAAGLSLPSETSPILRAEDPVPFPPSPKPIVDGSKTT 19602
Qy 156 ----- 155
Db 19603 ITIAWVKPLFDGAPITGYTVEYKKSDDTDWKTISQSLRGTEYTSGLTTGAEYVFRVKS 19662

Qy 156 ----- 155
Db 19663 VNKVGASDPSSDPQIAKEREPEPLFDIDSEMRKTLIVKAGASFTMTVPRGRPVNVL 19722
Qy 156 ----- 155
Db 19723 WSKPDTDLTRAYVDTTDSRTSLTIENANRNDGKYTLTIQNVLSAASLTLLVKKVLDTPG 19782
Qy 156 ----- 155
Db 19783 PPTNITQDVTKESAVLSWDVPENDGGAPVKNYHIEKREASKAWSVTNNCRLSYKVT 19842
Qy 156 ----- 155
Db 19843 NLOEGAIYFRVSGENEFVGIPAEWKEGVKITEKPSPEKLGVTISKDSVSLTWLKPE 19902
Qy 156 ----- 155
Db 19903 HDGSRIVHYVEALEKGQKNWYKCAVAKSTHHVWVSGLENSEYFFRFAENQAGLS DPR 19962
Qy 156 ----- 155
Db 19963 ELLLPVLIKEQLEPPEIDMKNFPSHTVYVRAGSNLKVDIPISGKPLPKVTLSDRGVPLKA 20022
Qy 156 ----- 155
Db 20023 TMRPNTETAENLTINLKESVTADAGRYEBITAANSSGTTKAFINIVVLDPRPGPTGPVVI 20082
Qy 156 ----- 155
Db 20083 SDITEBSVTLKWBPKYDGGSQVTNYILLKRETSTAVTVEVSATVARTMMKVKLTGTEE 20142
Qy 156 ----- 155
Db 20143 YQPRIKAENRFGISDHIDSACVTVKLPYTPPGPPSTPWNTVNTRESTITGVHHEPVSNGGS 20202
Qy 156 ----- 155
Db 20203 AVVGYLEMKDRNSILWQANKLIVIRTHPKVTTISAGLIYEFVRYAENAGVKPSHPS 20262
Qy 156 -----ELTKNOVSL----- 164
 :::|:|:|
Db 20263 EPVLADACEPPRNVRTIDSKNSVLSWQOPAFDGGSKITGYIVERRDLDPGRWTKASF 20322
Qy 165 -----TCL----- 167
 ||:
Db 20323 TNVTETQFTISGLTQNSQYEFVRFARNAGVSGISNPSEVVGPIITCIDSYGGPVIDLPLEYT 20382
Qy 168 ----- 167
Db 20383 EVVKYRAGTSVKLRAGISGKPAPTIEWYKDDKELOTNALVCVENTD LASILIKDADRLN 20442
Qy 168 ----- 167
Db 20443 SGCYELKLRNAMASATIRVQILDKPGPPGPIEFKVTVAEKITLLWRPPADGGAKIT 20502
Qy 168 ----- 167
Db 20503 HYIVEKRETSRVVWVMSVSEHLEECIITTTKIIKGNEYIFRVRVAVNKYGIGEPLESDSVA 20562
Qy 168 ----- 167
Db 20563 KNAFVTPGPGPIEVTKITKNSMTVYVSRPIADGGSDISGYFLEKRDKSLGWFKVLKET 20622
Qy 168 ----- 167
Db 20623 IRDTROKVTGLTENS DYQYRVCAVNAAGOGFSEPFYKAADPIDPGPPAKIRIADST 20682
Qy 168 ----- 167
Db 20683 KSSITLGWSKPVYDGGSAVTGYVVEIROGEEBWTTVSTKGEVTRTEYVYVSNLKPUNYY 20742

QY 168 ----- 167
Db 20743 FRSAVACAGGEP IEMNEPQAKIDLEAPEIDLDVALRTSVIAKAGEDVQLIPFKGRP 20802
QY 168 ----- 167
Db 20803 PPTVTRKDEKNLGS DARYSIENWTDSSLLTIPQVTRNDTGKXILTIENGVGEPKSTVS 20862
QY 168 ----- 167
Db 20863 VKVLDTPAACOKLOVHVSRGTVTLLWDPPLIDGGSPINNVIEKRDATAKRTWSVSHKC 20922
QY 168 ----- 167
Db 20923 SSTSKLIDLSEKTPPFPRVLAENEIGIGPCETTEPVKAAEPAPIRDLMSKDSTKTSV 20982
QY 168 ----- 167
Db 20983 ILSWKDPDPGGSVITEYVVERKKGQWTHAGISKTCEIEVSQLEQSVLEFRVPAKN 21042
QY 168 ----- 167
Db 21043 EKLSDPVTIGITVRELIIITPEVDSIDIPCAQVTVRIGHNVHLELDYKGPKPSPISWLK 21102
QY 168 ----- 167
Db 21103 DGLPKESFVRFSKTENKITLSIKNAKKEHGKYTVILDNAVCRIVAPITVITLGPSPK 21162
QY 168 ----- 167
Db 21163 PKGPIRDEIKADSVILSDVPEDNGGGEITCYSEKRETSQTNWKMVC SVARTTEKVP 21222
QY 168 ----- 167
Db 21223 NLVKDAEQFRVRAENRYGVSQPLVSSIIIVAKHQFRIPGPGKPVIVNVTS DGMSLTWDA 21282
QY 168 ----- 172
Db 21283 PVDGGSEVTGFHVKEKERNLSILWQVNTSPISGREYRATGLVEGLDYQFRVYAENSAGL 21342
QY 173 --PSD----- 175
Db 21343 SSPSDPSKFTLAVSPVDPGTPDYIDVRETITLKNPPLRDGSGKIVGYSIEKROGNER 21402
QY 176 ----- 175
Db 21403 WVRNFTDVSECQYVTGLSPGDRYEFRIIARNNAVGTISPSPSSGIIIMTRDENVPPIVE 21462
QY 176 ----- 175
Db 21463 FGPEYFDGLIIKSGESLRKALVQGRPVPRVTFWKDGVETIEKRMNMEITNVLGSTSLFVR 21522
QY 176 ----- 175
Db 21523 DATRDHRGVYVEAKNAGSAKAEIKVKQDTPGVVGPVIRFTNITGERMTLWMDAPLND 21582
QY 176 ----- 175
Db 21583 GCAPITHYIIEKRETSRLAWALIEDKCEAOSYTAIKLINGNEYQFRVSAVNKFGVGRPLD 21642
QY 176 ----- 175
Db 21643 SDPVAIQYTVDPAGIPEPSNITGNSITLTWARPESDGGSEIQOYILERREKKSTRWV 21702
QY 176 ----- 175
Db 21703 KVISKRPISETRFKVTGLTEGNEYEFHVMAENAAAGVPASGISRLIKREBPVNP GPPTV 21762
QY 176 ----- 180
Db 21763 VKVTDTSKTVSLENSKPVFDGMEIIGYIIECMCKTDLGDWHKVNACVKTYYTVDLQ 21822
QY 181 ----- 180

Db 21823 AGEYKFRVSAINGAGKGDSCVETGTIKAVDRLTAPELDIDANFKQTHVVRAGASIRLFI 21882
QY 181 ----- 183
Db 21883 AYOGRRPTAVWSKPDNSLSRADIHITDSFSTLTVCNCRNDAGKYTLTVENNSGSKSI 21942
QY 184 ----- 183
Db 21943 TFTVKVLDTPCPPGPITFKDVTGRSATLMDAPILLDGGARLHHYVVEKREASRSWQVIS 22002
QY 184 ----- 183
Db 22003 EKCTRIQIFKVNDLAEGVPYIFRVSANVEYGVGEPEMPEPIVATEQAPPPRLDWDTSK 22062
QY 184 ----- 183
Db 22063 SSAVLAWLKPDPHDGGSRTGYLLEMROKGS D LWVEAGHTKOLTFTVERLVEKTEYEFVRK 22122
QY 184 ----- 183
Db 22123 AKNDAGYSEPREAPSSVIIKEPQIEPTADLTGITNQLITCKAGSPFTIDVPISGRPAKV 22182
QY 184 ----- 183
Db 22183 TWKLEMLRKETDRVSITTTKDRITLTWKDSMRGDSGRYFLLTENTAGTKFTSVTVVVG 22242
QY 184 ----- 186
Db 22243 RGPVTPGPIEVSSVSAESCVLWSGPEKDGSGTEITNYIVEKRESGTTAWQLVNSSVRQTQ 22302
QY 187 ----- 186
Db 22303 IKVTHLTKYMEYFRVSSSENRFVSKPLESAPIIAEHFPVPSPAPTRPEVHVHVSANMSI 22362
QY 187 ----- 193
Db 22363 RWEEPYHDGSGKIIGYWEKKERTILWVKENKVPCLCNKVTGLVEGLEVOFTVALN 22422
QY 194 ----- 193
Db 22423 AAGVSKASEASRPIMAQNPDVAPGRPEVDVTRSTVSLINSAPAYDGGKVVGYIIEKRP 22482
QY 194 ----- 193
Db 22483 VSEVGDGRWLKCNYYIVSDNFFVTALSEGDTYEFRLAKNAAGVISKSESTGPVTCRD 22542
QY 194 ---pp----- 195
Db 22543 EYAPPKAELDARLHGLVTRAGSDLVLDAAVGGKPEPKIIWTGDKDELDCENKSLQYT 22602
QY 196 ----- 201
Db 22603 GKRATAVIKFCDRSDSGKYTLTVKNASGTRAVSYMVVLDSPGCGKITVSRVTOEKCTL 22662
QY 202 ----- 201
Db 22663 AWSLPQDGGAEITHYIVERRETSRLNWVIVEGECPTLSYVVTRLIKNNEYIFRVRVANK 22722
QY 202 ----- 205
Db 22723 YGPGVPVESEPIVARNSTIPSPGPIPEVGTGKEHIIQWTKPESDGGNEISNVLVDKR 22782
QY 206 ----- 205
Db 22783 EKESLRWTRVNKYVYVDTRLKVTSLMEGCDYQFRVTAAGNAGSEPSERSNFISCREPS 22842
QY 206 ----- 205
Db 22843 YTPGPPSAPRVVDTTKHSISLAWTKPMVDGGTDIVGVYLEMOEKDQWVRVHTNATIRN 22902
QY 206 ----- 205

Db 22903 TEFTVPLKMGOKYSRVAANVNGMSEYSEIAETIEPVERIEIPDELEADDLKKTVTIR 22962
QY 206 -----YSLTVDK----- 213
Db 22963 AGASLRLVSVSGRPPVITWSKQIDLASRALIIDTESYLLIVDKVNRDAGKYTIEA 23022
QY 214 ----- 213
Db 23023 ENQSGKKSATVLVKVYDTGCPSPVKVKSRSDSVTITWEIPTIDGGAPINNVIVEKREA 23082
QY 214 ----- 213
Db 23083 AMRAFKVTTKCKSLYRISGLVEGTMHYFRVLPENIYGIGEPCESTDALVSEVPLPA 23142
QY 214 ----- 213
Db 23143 KLEVVDVTKSTVLAWEKPLYDGGSLRTGYVLEACKAGTERWKMVVTLKPTVLEHTVTSI 23202
QY 214 ----- 213
Db 23203 NEGEQYLFRIQAQNEKGVSERETVAVTVQDLRLVPTIDLSTMPKTIHVPAGRPVELV 23262
QY 214 ----- 213
Db 23263 IPIAGRPPPAASWFFAGSKLRESERVTVETHKVAKLTITRETTIDTGYTILELKNVTGT 23322
QY 214 ----- 213
Db 23323 TSETIKVIILDKPGPTGPIKIDEIDATISITISWPELDDGAPLSGVVVEQDAHRPQW 23382
QY 214 ----- 213
Db 23383 LPVSESVRSTRTEKFTRLTEGNEVYFVAAATNRFGIGSYLQSEVIECRSSIRIPGPETIQ 23442
QY 214 ----- 213
Db 23443 IFDVSRDGMTLWYPPEDDGGSOVTGYIVERKEVRADRVVRVKNVPTWTRYSTGLTEG 23502
QY 214 ----- 213
Db 23503 LEVEHRTAINARGSKPSRPKPIVAMPDIAPPKQPQPRVTDTRTSVSLAWSPEDE 23562
QY 214 ----- 213
Db 23563 GGSKVTYLIEMQKVDQHEWTKCNTPTKIREYTLTHLPOGAERYFRVLACNAGGPGEPA 23622
QY 214 -----SRWQOQ----- 219
Db 23623 EVPGTVKVTEMLEYDPYELDERYQEGIFVRQGGVIRLTPIPIKGPFPICKWTKEGODISK 23682
QY 220 ----- 219
Db 23683 RMIATSEHTELVIEADRGDSGYDVLVLEKNCKGKAYYIKVRIGSPNSPEGPLEYDD 23742
QY 220 ----- 219
Db 23743 IQVRVRSWRPPADGGADILGYILERREVPKAAWTTIDSRVGTSLVVKLGLENVEYH 23802
QY 220 ----- 219
Db 23803 FRVSAENQGISKPLKSEEPVTPKTPLPNPPEPPSNPPEVLDVTKSVSLSWSRPKDDGGS 23862
QY 220 ----- 219
Db 23863 RVTGYIERKETDKVVRHNKTQITTMVTVTGLVPDAEQFRIAQNDVGLSETSPAS 23922
QY 220 ----- 219
Db 23923 EPVCKDPDKPSQCELEILSISKDSVTLOWEKEPCDGGKEILGWVVEYRQSGDSAWKK 23982
QY 220 ----- 219
Db 23983 SNKERIKDKQFTIGLLEATEYFRVFAENETGLSRPRRTAMSIKTKLTSGEAPGIRKEM 24042

QY 220 ----- 219
Db 24043 KDVTYTKLGEAAQLSCQIVGRPLPDIKWYRFGKELIOSRKYKMSSDGRTHTLVTMTTEQED 24102
QY 220 ----- 219
Db 24103 EGVYTCIATNEVEGETSSKLLLLQATPQHPGYPLKEYYAGVSTLRLHVMYIGRPVPA 24162
QY 220 -----NVF----- 222
Db 24163 MTWFHGQKLLQNSENITIENTHYTHLVKMNQVORKTHAGKYKVLQSNVFGTVDAILDVEI 24222
QY 223 ----- 222
Db 24223 QDKPKPTGPVIEALLKNSAVISMKPPADGGSWITNVVEKCEAKEGAEMQLVSSAIS 24282
QY 223 ----- 222
Db 24283 VTTCRIVNLTENAGYFRVSAQNTFGISDPLEVSSVYIIKSPFKPGAKPTITAVTKD 24342
QY 223 SCVV----- 226
Db 24343 SCVWAWKPPASDGGAKIRNYILEKREKKONKWIISVTTEIRETVFSVKNLIEGLEYEFRV 24402
QY 227 ----- 226
Db 24403 KCENLGGESEWSEISEPITPKSDVPIQAPHKEELNLRVYQSNATLVCKVTGHPKPIV 24462
QY 227 ----- 226
Db 24463 KMYROGKEIIADGLKYRIOEFKGYHQLIIASVTDDATVYQVRATNOGGSVSGTASLEV 24522
QY 227 ----- 226
Db 24523 EYPAKIHLPKTLLEGMAVHALRGEVVISIKPFGKPPDVITQKGQDLDNNNGHYQVIT 24582
QY 227 ----- 226
Db 24583 RSFTSLVFPNGVERKDAGFYVVCANRFGIDQKTVELDVADVPDPPRGYKVSASRDSVN 24642
QY 227 ----- 226
Db 24643 LTWTEPASDGSKITNYIVEKCATTAEERLVRGOARETRYVINLFGKTSYQFRVIAENK 24702
QY 227 ----- 226
Db 24703 FGLSPSEPSEPTITKEDKTRAMNYDEEVDETREVSMTKASHSTKELYEKYMAEDLGR 24762
QY 227 -----MHEA----- 230
Db 24763 GEGIVHRCVETSSKTYNAKFVKVKGTDQVLVKKEISILNIAHRNHLHLHESFESMEE 24822
QY 231 -----LHNH----- 234
Db 24823 LVNIFEFGIDIFERINTSAFELNEREIVSVYHVQCEALQFLHSHNIGHFDIRPENIY 24882
QY 235 -----YT----- 236
Db 24883 QTRRSSTIKIIEFGOARQLKPGDNFRLLFTAPEYYAPEVHQHDVVSVSTATDMNSLGLTVV 24942
QY 235 ----- 236
Db 24943 LLSGINPFLAETNQIIEINIMNAEYTFDEAFKEISIEAMDFVDRLLLVKERSRMTASEA 25002
QY 237 -----QKSLSL- 243
Db 25003 LQHPWLKOKIERVSTKVIRTLKHRRYYHTLIKKDLNMVYSAARISCGGAIRSQKGVSAK 25062
QY 244 ----- 243
Db 25063 VKVASIEIGPVSQIMHAVGEEGGHYVKYVCKIENYDQSTQVWTYFGVROLENSEKEYEITY 25122

```

QY 244 ----- 243
Db 25123 EDGVAILVVKDITKLDGTYRCKVNDYGEDSSYAEFLVKGREVVYDYCRRTMKIKRR 25182
QY 244 ----- 243
Db 25183 TDTMLLERPFTPLYNKNTAYGVENRFGVTITVHPPEHVTWYKSGQIKPGDNKKY 25242
QY 244 ----- 243
Db 25243 TPESDKGLYQLTINSVTTDDDAEYTVVARNKYGEDSKAKLVTLHPPTDSTLRPMFKR 25302
QY 244 ----- 243
Db 25303 LLANAEQEQSQVCEIRVSGIPPTLKWEKDGQPLSLGNPTIEIHEDLDYVALHIDTFL 25362
QY 244 ----- 243
Db 25363 PEDTGYRVTAATAGSTSCQAHQVRLRYKKQEFKSEHERHVQKQIDKTLRMAEIL 25422
QY 244 ----- 243
Db 25423 SGTESVPLTQVAKAEALREAAVLYKPAVSTKVKGFRLEIEKKERKLMPDYDPEPRK 25482
QY 244 ----- PGK 246
Db 25483 YKQTTIEEDQRIKQFVPMDSMKWYKIRQYEMPGK 25518
[1]
[2]
RESULT 15
Q8WZB3
ID Q8WZB3 PRELIMINARY; PRT: 26926 AA.
AC Q8WZB3;
DT 01-NAR-2002 (TremBLrel. 20, Created)
DT 01-WAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE N2B-titin isoform.
GN TTN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20309627; PubMed=10850961;
RA Freiburg A., Trombitas K., Hell W., Cazorla O., Fougereousse F.,
RA Centner T., Kolmerer B., Witt C., Beckmann J.S., Gregorio C.C.,
RA Granzier H., Labelt S.;
RT "Series of exon-skipping events in the elastic spring region of titin
RL as the structural basis for myofibrillar elastic diversity.";
RL Circ. Res. 86:1114-1121(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21573839; PubMed=11717165;
RA Bang M.L., Centner T., Fornoff F., Geach A.J., Gotthardt M.,
RA McNabb M., Witt C.C., Labelt D., Gregorio C.C., Granzier H.,
RA Labelt S.;
RT "The complete gene sequence of titin, expression of an unusual -700
kDa titin isoform and its interaction with obscurin identify a novel
RT Z-line to I-band linking system.";
RL Circ. Res. 89:1065-1072(2001).
DR EMBL; AJ277892; CAD12455.1;
DR InterPro; IPR000282; Cytochrome receptor_2.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR000577; FGGY_kin.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003600; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR002016; Peroxidase.
DR InterPro; IPR002290; Ser_thr_pkinase.

```

```

InterPro; IPR001245; Tyr_pkinase.
Pfam; PF00041; fn3; 132.
Pfam; PF00047; ig; 91.
Pfam; PF00069; pkinase; 1.
ProDom; PD000001; Euk_pkinase; 1.
SMART; SM00060; FN3; 132.
SMART; SM00409; IG; 111.
SMART; SM00408; IGc2; 93.
SMART; SM00410; IG-like; 3.
SMART; SM00220; S_TKc; 1.
SMART; SM00219; TyrKc; 1.
PROSITE; PS00933; FGGY_KINASES_1; UNKNOWN_1.
PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
PROSITE; PS00290; IG_MHC; UNKNOWN_1.
PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
SQ SEQUENCE 26926 AA; 2993285 MW; 169AB42637A7C1FB CRC64;

Query Match 34.7%; Score 463; DB 4; Length 26926;
Best Local Similarity 0.7%; Pred. No. 0.00021;
Matches 168; Conservative 43; Mismatches 34; Indels 23731; Gaps 53;

QY 2 WLK----- 4
Db 1543 WLKSDIIIVPHKPKIRIEGTKEAALKIDSTVSQDSAMVTATAINKAGRDITTRCKVNVE 1602
QY 5 -----AFYK----- 9
Db 1603 VEFAPEPERKLIIPRGTYRAKEIAAPELEPLHLRYGOEQWEGDLYDKEKQKPFKK 1662
QY 10 ----- 9
Db 1663 LLSRLKRFPAHFECRLTPIGPTMVVWELHDGCKLEAANRLMINERGCYSLDYGVAY 1722
QY 10 ----- 9
Db 1723 SRDSGIITCRATNKYGTDTHTSATLIVKDEKSLVEESQLEPGRKGLQRIEELERMAHEGAL 1782
QY 10 ----- 9
Db 1783 TGVTTDOKEQKQPDIVLYPEPVVLEGETARFCRVTPQPKVNWYLNGLIRKSKRFR 1842
QY 10 -----VAE----- 12
Db 1843 VRYDGHLYLDIVDCKSYDTGEVKVTAENEGVIEHKVLEIQOERDFRSVLRRAPERPE 1902
QY 13 -----KLKEA----- 17
Db 1903 FHVHEPGKQFEVQKVDPRVDTTETKEVVKLKAERITHEKVPSESELRSKFKRTEEG 1962
QY 18 ----- 17
Db 1963 YYEAITAVELKSRKKDESYEELLRTKTKDELLHWTKELEKALAEKGKITIPTFKPKD 2022
QY 18 ----- 17
Db 2023 IELSPSEAPKIFERIQSTQVGGSDAHPFRVVGPDCEWYKNGVKIERSDRIYWW 2082
QY 18 ----- 17
Db 2083 PEDNVCVELVTDVAEDSASIMVKAINIAGETSSHAFLLVQAKLITFTQELQDVVAKER 2142
QY 18 ----- 17
Db 2143 DTMATFECETSEPFVKVKWKYKDGMEVHEGDKYRMHSDRKVHFLSILTIPTSDAEDYSCVL 2202
QY 18 ----- 17
Db 2203 VEDENVKTTAKLIVEGAVVEFVKELQDIEVPESYSGELECIIVSPENIEGKWHNDVELKS 2262
QY 18 ----- 17

```

Db 2263 NGKYTITRRGRQNLTVKDVTKEDQGEYSFVIDGKKTCTCKLMKPRPTAILOGLSDQKVC 2322
Qy 18 ----- 17
Db 2323 EGDIVQLEVKYSLESVEGVWMDQGEVQPSDRVHIVIDKQSHMLLIEDMTKEDAGNYSFT 2382
Qy 18 ----- 17
Db 2383 IPALGLSTGRVSVYSDVITPLKDVNVIETKAVLECKVSPDVTSVKWLNDQIKPD 2442
Qy 18 ----- 17
Db 2443 DRVQAIKGTOKORLVINRTHASDEGPYKLIIVGRVETNCLNVEKIKIIRGLRDLTCTETQ 2502
Qy 18 ----- FMDK----- 21
Db 2503 NVFVELSHSGIDVLNFKDKEIKPSSKYKIEAHGKIYKLTVLNMMKDDGKYTFYAGE 2562
Qy 22 ----- 21
Db 2563 NITSGKLVAGGAISKPLTDQVAESQEAVFCEVANPDSKGWLRDGLPLTNIRSE 2622
Qy 22 ----- 21
Db 2623 SDGHKRRLLIAATKLDIDIGEITYKVATSKTSAKLKVEAVKIKTKLNLTVTETQDAVFTV 2682
Qy 22 ----- 21
Db 2683 ELTHPNVKGWIKNGVLESNEKYAISVKGTIYSLRIKNCIIVDESIVGPRGLGASA 2742
Qy 22 ----- 21
Db 2743 RLHVETVKI IKKPKDVTALENATVAFVSVSHDTPVKWFKHVEIKPDKHRLVSEKV 2802
Qy 22 ----- 21
Db 2803 HKMLQNI SPDAGEYTA VVGQLECKAKLFVETLHITKMKNIEVPETKTASFCEVSHF 2862
Qy 22 ----- 21
Db 2863 NVPSMLKNGVEIEMSEKFIWQGLHQLIIMNTSTEDSAEYFVCGNDQVSATLVTVP 2922
Qy 22 -----THT----- 24
Db 2923 IMITSMKLDINAEKDTITFEVTNVEGISYKWLKNGVEIKSTDKQCMRTKKLTHSLNIR 2982
Qy 25 ----- 24
Db 2983 NVHFGDAADYTFVAGKATSTATLYVEARHIEFRKHIDIKVLEKKRAMFECEVSEPDITV 3042
Qy 25 ----- 24
Db 3043 QWMKDDQELQITDRIKIQEKYVHRLIIPSTRMSDAGKYTVVAGNVSTAKLFEVGRDVR 3102
Qy 25 ----- 24
Db 3103 IRSIKKEVQVIEKQRAVVEFEVNEDDVAHWYKDGIEINFQVQERHKYVVVERIRHMFIS 3162
Qy 25 ----- 24
Db 3163 ETROSDAGEYTFVAGRNRSSVTLXVNAPEPPQVLQELQVTVVQSGKPARFCAVISGRPQP 3222
Qy 25 ----- 24
Db 3223 KISWYKEQLLSTGFKCFLHDGOEYTLILLIEAFPDAAYTCEAKNDYGVATTSASLSV 3282
Qy 25 ----- 24
Db 3283 EVPEVSPQEMPVYPAITPLQDTVTISEGPQARFCRVSGTDLKVSWSYKDKKIRPSR 3342
Qy 25 ----- 24
Db 3343 FFRMTQFEDTYOLEIAEYAPEDEGTYTFVASNAGVQVSSSTANLSLEAPESILHERIEQEI 3402

Qy 25 ----- 24
Db 3403 EMENKEFSSSFLSAAEEGLHSAELQLSKINETLELLSEPVYSTKFDSEKEGTGPIFIKE 3462
Qy 25 ----- 24
Db 3463 VSNADISMGDVATLSVTVIGIPKPKIQWFFNGVLLTPSADYKFFVDDGDDHSLIILFTKLE 3522
Qy 25 ----- -cpp----- 27
Db 3523 DEGEYTCMASNDYKGTCTCSAYLKNKSGEGHKDTETESAVAKSLEKLGCGPCPPHFLKELK 3582
Qy 28 ----- 27
Db 3583 PIRCAOGLPAIFEYTVVGEPAPTVTFWKENQOLCTSVYTTIHNPNGSTGTFVNDPQRED 3642
Qy 28 -----CPAPELLG----- -GPSV--- 39
Db 3643 SGLYICKAENMLGESTCAAELLVLEDDMTDTPCKAKSTPEAPEDFPQTLKGPAAVEAL 3702
Qy 40 ----- 39
Db 3703 DSEOEIATFVKDTILKAALITEENOQLSYEHIAKANELUSSQLPLGAQLQILSLODKLTP 3762
Qy 40 -----FL----- 41
Db 3763 ESTREFLCINGSIHFPOLKEPSPNLQLOIVOSOKTFSKEGILMPEEPETOAVLSDEKIF 3822
Qy 42 ----- 41
Db 3823 PSAMSIQINSLTVEPLKTLAEPEGNYPOSSIEPPMHSLTSVAEEVLSPEKTKVSDTN 3882
Qy 42 ----- 41
Db 3883 REQVTLQKQEAQASALILSQSLAEHVESLOSPOVMISQVNYEPLVPSEHSTEGGKILI 3942
Qy 42 -----FP----- 43
Db 3943 ESANPLENAGQDSAVRIEESKSLRFPPLALEEKQVLLKEHSDNVVMPDQIIESKREPA 4002
Qy 44 ----- 43
Db 4003 IKKQVEQGRDLLSKESLLSGIPEEQRLNLKIQICRALQAAVASEQPGFSEWLRNIEKV 4062
Qy 44 ----- 43
Db 4063 EVEAVNITQEPRHIMCWMYLVTSKASVTEEVTTIIEDVDPQMANLKMELRDALCAIIEEI 4122
Qy 44 ----- 43
Db 4123 DILTAEGPRIQOGAKTSLQEEMDSFSGSKVEPITEPEVESKYLISPEVSYFNVQSRVK 4182
Qy 44 ----- 43
Db 4183 YLDATPVTKGVASAVVSDQESLKPSEKESSESSESTEEVATVKIQEAEGGFIKBDG 4242
Qy 44 ----- 43
Db 4243 PMIHPTPLVDTVSEBGDIVHLTTSITNAKEVNWYFENKLVPSDEKFKCLQDQNTYTLVIDK 4302
Qy 44 ----- 43
Db 4303 VNTEHQGEYVCEALNDSGKTATSAKLTVVKRAAPVIKRIEPLLEVALGHLAKFTCEIQS 4362
Qy 44 ----- 43
Db 4363 APNVRFOWKAGREIYESDKSIRSKYISSLEILRTQVDCGEYTCCKASNEYSVSCA 4422
Qy 44 -----PKPKDTLMI----- 52
Db 4423 TLTVTVPGEKKVKRLLLPKRPKEPEVVLKSVLRKRPEEPEKPKLEKKVPAVPE 4482

QY 53 ----- 52
Db 4483 PPPPKPVEEVTVTKRERK IPEPTKVPEIKPAIPLPAPEPKPPEAEVKTIKPPPEP 4542
QY 53 ----- SRTP----- 56
Db 4543 EPTPIAAPVTVPVVGGKAAKAPKEAAKPKGPIKGVPKKTSPSIEARRKLRPGSGEK 4602
QY 57 ----- 56
Db 4603 PPDEAPFTYQKAVPLKFKVKEIKDIILTESEFVGSSAIFECIVSPSTAITTMMKDGSNIR 4662
QY 57 ----- EVTCV----- 61
Db 4663 ESPKHFRIADGDKRLHIIDVQLSDAGEYTCVLRGNKEKTSFPAKLVEELPVRFVKTLE 4722
QY 62 ----- III----- 61
Db 4723 EEVTVVKGQPLXLSCELNKERDYYVRKDGKIYVEKPGRIYVGVIGLRAITINDADDTDA 4782
QY 62 ----- 61
Db 4783 GTYTVTENANNLECSVCVVVEVIRDLVKPIRDQHVPKGTAIFACDIAKDTPNIKWF 4842
QY 62 ----- 61
Db 4843 KGYDEIPAEPNDKTEILRDGNHLYKIKNAMPEDIAEYAVEIBGKRYPAKLTJGEREVEL 4902
QY 62 ----- 61
Db 4903 LKPIEDVTIYKESAFDAISEADIPGOWKLGELLRSPPTCEIKAEGGKRFLLRKVK 4962
QY 62 ----- 61
Db 4963 LDQAGEVLYQALNAITAILTVKEIELDFAVPLKDVTVPERRQARFECVLTREANVIWSK 5022
QY 62 ----- 61
Db 5023 GPDIIKSSDKFDIIADGKKHILVINDSQFDEGVYTAEVGKTSARLFTVGIKLFMSF 5082
QY 62 ----- VDVSH----- 68
Db 5083 LEDQTVKEGETATFYCGLSHEKHMVYVFNDAKLHTSRTVLISSEKTHKLEMKEVTLDD 5142
QY 69 ----- 68
Db 5143 ISQIKAQVKELSTAQIKVLEADPYFTVKLHDKTAYEKDEIILKCEVSKDVPVKWFKDGE 5202
QY 69 ----- 68
Db 5203 EIVSPKYSIKADGLRILKIKKADLKDKGEYVDCGTDKTKANVTVEARLIKVEKPLYG 5262
QY 69 ----- 68
Db 5263 VEVFVGETAHEIELSEPDVHGOWKLGQPLTASPDCEIIEDGKKHILHNCOLQMTGE 5322
QY 69 ----- 68
Db 5323 VSFQAAKSAANKVKELPLFIITPLSDVKVFEKDEAKFECEVSREPKTFRWLKGTQEI 5382
QY 69 ----- 68
Db 5383 TGDDRFELIKGTGKHSWIKSAAFEDAKEYMFEADKHTSGKLIIEGIRLKFUTPLKDVT 5442
QY 69 ----- 68
Db 5443 AKEKSAVFTVELSHDNIYVKFNQDRLHTTRSVSMQDEGKTHSITFKDLSIDDTSQIR 5502
QY 69 ----- DP----- 70
Db 5503 VEAMGSSSAKLTVLEGGDPTFTGKLODYGVEKDEVILOCEISKADAPVKWFKDGKEIRP 5562
QY 71 ----- 70

Db 5563 SKNAVIKADGKKRMLILKKALKSDIGQYTCDCGTDKTSGLKDIEDREIKLVRPLHSVEVM 5622
QY 71 ----- 70
Db 5623 ETETARFETEISEDDIHANWKLKGEALLQPPDCEIKEEGKIHSILVHNCRLDQTGGVDFQ 5682
QY 71 ----- 70
Db 5683 AANVSSAHLRVKPRVIGLRLPLKDVTVTAGETATFDCELSYEDIPVENVILKGGKLEPSD 5742
QY 71 ----- 70
Db 5743 KVVPRSEGVHTLTLRDVKLEDAGEVQLTAKDFKTHANLFVKBPPEFTKPLEDQTVEEG 5802
QY 71 ----- 70
Db 5803 ATAVLECEVSRENKVKWFKNGTEILKSKYIEIADGRVRKLVIHDCPTEDIKTYTODAK 5862
QY 71 ----- 70
Db 5863 DFKTSCNLNVVPPHVEFLRPLTDLQVREKEMARFECELSRENKVKWFKDGAELKGGKY 5922
QY 71 ----- 70
Db 5923 DIISKGAVRILVINKLLDDEAEVSCVETARTSGMLTVLEEEAVFTKNLANIEVSETDT 5982
QY 71 ----- 70
Db 5983 IKLVCEVSKPGAENVYWKGDDEIITGRYEILTEGRRILVIONAHLEDAGNINCRLPSS 6042
QY 71 ----- 70
Db 6043 RTDGKVKVHELAAEFISKPONLEILEGEKAEFVCSISKESFPVQWKRDDKTLESBGKYDV 6102
QY 71 ----- EVKF 74
Db 6103 IADGKKRVLVKDATLQDMGTYYVMGAARAAHLTVIEKLRIVVPLKDFRVEQEQVWF 6162
QY 75 N----- III----- 75
Db 6163 NCEVNTEGAKAKWFRNEEAIFDSSKYIILQKDLVYTLRIRIDAHLDQOANVSLTNHRGE 6222
QY 76 ----- 75
Db 6223 NVKSAANLIVEEDLRIVERPLKDIEITMEKKSVTFWCKVNRNLVTLKWTKNGBEEVFPDNRV 6282
QY 76 ----- 75
Db 6283 SYRVDKYKHLMTIKDGGFPDEGEYIVTAGQKSVAEELLIIEAPTEFVEHLEDQTVTEFDD 6342
QY 76 ----- 75
Db 6343 AVFSCOLSREKANVKWYRNGREIKEGKKYFEKDGSTHRLIIKDCRLDDECEYACGVEDR 6402
QY 76 ----- 75
Db 6403 KSRARLFVEEIPVEIIRPPODILEAPGADVVFVLAELNKKDKVEVQWLRNNVWVVGDKHQM 6462
QY 76 ----- 75
Db 6463 MSEGKTHRLQICDIKPRDQGEYRFAKDKARAKELAAAPKIKTADQLDVLVDVGKPLTM 6522
QY 76 ----- 75
Db 6523 VVPYDAYKAAEWFKENEPSTKTIDTTAEQTSFRILEAKGDKGRYKIVLQNKHGKAE 6582
QY 76 ----- 75
Db 6583 GFINKLVIDVPGVPRNLEVTFETFDGEVSLAWEEPILDGSGKIIGYVYVRRERRDIKRTWVLA 6642
QY 76 ----- 75

Db 6643 TDRASCEFTVTGLQGGVEYLFVRSAARNRVTGEPVETDNPVEARSKYDVPGPLNVTI 6702
Qy 76 ----- 75
Db 6703 TDVNRFGVSLTWEPPEYDGGABITNVYIELRDKTSIRWDTAMTVRAEDLSATVTDVWEGQ 6762
Qy 76 ----- 75
Db 6763 EYSFRVRAQRIGVGKPSAATPFVKVADPIERPSPVNLTSDDQTSQSSVOLKWEPLKDG 6822
Qy 76 ----- 75
Db 6823 GSPILGIIERCEEKGNWIRCNMKLVPETYKVTGLEKGNKLYRVSAENKAGVDPSE 6882
Qy 76 ----- 75
Db 6883 ILGPLTADDAFVEPTMDLSAFKDLGVIVPNPITILVPSTGYPRPTATWCEGDKVLETGD 6942
Qy 76 ----- 75
Db 6943 RVKMKTL SAYAELVISPSERSDKGIYTLKLENRVKTSISEIDVNVIA RPSAPKELKFGDI 7002
Qy 76 ----- 75
Db 7003 TKDSVHLTWEPDDGGSPLTGYVVEKREVSRTWTVMDFVTDLEFTVPDLVQKGYLF 7062
Qy 76 ----- 75
Db 7063 KVCARNKCGGEPAYVDEPVNMSTPATVPDPENVKWRDRTANSIFLTWDPKNDGSR I 7122
Qy 76 ----- 75
Db 7123 KGYIVERCPGSDKKWACGEPVAETKMEVTGLEEGKWAYRVKALNRQASKPSRPTBEI 7182
Qy 76 ----- 75
Db 7183 QAVDTQEAPEIFLDVKLLAGLTVKAGTKIELPATVTKRPEKITWTKADMILKODKRITI 7242
Qy 76 ----- 75
Db 7243 ENVPKSTVTIVDSKRSDTGYIIEAVNVCGRATAVVEVNVLDKPGPPAAFDITDVTNES 7302
Qy 76 ----- 75
Db 7303 CILLTNPPRDDGGSKITNVVERRATDSEVWHKLSSTVKTDNFKATKLIPNKEYIFVAA 7362
Qy 76 ----- 75
Db 7363 ENMYGVGEVQASPTAKYQDFDPGPPTRLEPSDITKDAVTLTWCEPDDGGSPITGYWV 7422
Qy 76 ----- 75
Db 7423 ERLDPDIDKWRNCNMPKYDTYRVKGLTNKKYFRVLAENLAGPGKPSKSTEPILIKD 7482
Qy 76 ----- 75
Db 7483 PIDPPWPGKPTKVDGKTSVRLNWKPDHGGAKIESVTIEMLKGTGDWVRVAEGVPT 7542
Qy 76 ----- 75
Db 7543 TOHLLPGLMEGOEYSFRVRAVNKAGESESPSDPVLCREKLYPPSPRMLVINITKNT 7602
Qy 76 ----- 75
Db 7603 ADLKWTPKEKGGSPITNVI VEKRDVRKGMQVTDITVKDTKCI VTPLTEGSLYVFRVAA 7662
Qy 76 ----- 75
Db 7663 ENAIGQSDYTEIDSVLAKDTFTTPGPPYALAVDVTKRHRVDLKWEPKNDGGRPIQRYV 7722
Qy 76 ----- 75
Db 7723 IEKKERLGRWVKAGTAGPCNFRVTDVIEGTVEQFQVRAENEAGVCHPSEPTEILSIE 7782

Qy 76 ----- 75
Db 7783 DPTSPSPPLDLHVTDAGRKKHIAIAWKPEKNGGSPITGYHVMCPVGT EKKMRVNSRPI 7842
Qy 76 ----- 75
Db 7843 KDLKFKEGVVVPDEKEYVLRVRVNAIGVSEPSEISENVVAKDPCKPTIDLETHDIIVI 7902
Qy 76 -----WYVDG----- 80
| : ||
Db 7903 EGKLSIPVFPRAVPVPTVSMHKDGKEVKASDRLTMKNDHISAHLEVPKSVRADAGIYTI 7962
Qy 81 ----- 80
Db 7963 TLENKLGSAATASINVKVIGLPGCPCKDIKASDITKSSCKLTWEPPEFDGGTPIILHYVLERR 8022
Qy 81 ----- 80
Db 8023 EAGRTYIPVMGSENKLSMTVKDLIPNGEYFFRVKAVNKVGGEYIELKNPVIAQDPKQP 8082
Qy 81 -----VEVHNA----- 86
|||||
Db 8083 PDPPVDVEVHNPTAEAMTITWKPPLYDGGSKIMGYIIEKIAKERWRKRCNEHLVPILTY 8142
Qy 87 ----- 86
Db 8143 TAKGLEEGKEYQFRVRAENAAGISEPSRATPPTKAVDPIDAPKVLIRTSLEVKGDEITAL 8202
Qy 87 ----- 86
Db 8203 DASISGSPYPTITWIKDENVIVPEIKKRAAPLVRRRKGEVQEEEPFVPLTLQRLSIDNS 8262
Qy 87 ----- 86
Db 8263 KKGESQLRVRDSL RPDHGLYMIKVENDHGIAPCTVSVLDTPGPINFPVEDIRKTSVL 8322
Qy 87 -----KTKP----- 90
|||||
Db 8323 CKWEPPLDDGGSEIINYITLEKKDKTKPQSEWIVTSTLRHCKYSVTKLIEGKEYLFRVA 8382
Qy 91 ----- 90
Db 8383 ENREGPGPCVSKPLVAKDPFGPDAPDKPIVEDVTSNSMLVKWNEPKDNGSPILGYWLE 8442
Qy 91 ----- 90
Db 8443 KREVNSTHMSRVNKSLLNALKANVDLLEGLTYVFRVCAENAGPGKPSPPDPKTAHDP 8502
Qy 91 ----- 90
Db 8503 ISPPGPIPRVTDTSSTTIELEWEPAPNGGGEIVGYFVDKQLVGTNEMSRCTEKMIVKR 8562
Qy 91 ----- 90
Db 8563 QYTVKEIREGADYKLRVSAVNAAGEPPGETQPVTVABEQPAPVAVELDVSVKGGIQUIMAG 8622
Qy 91 ----- 90
Db 8623 KTLRIPAVVTRGPVPTKVWTKEBGELDKDRVVIDNVGTKSELIIKDALRKDHGRYVITAT 8682
Qy 91 ----- 90
Db 8683 NSCGSKFAAARVEYDVPGPVLDLKPVTNRKMCLLNWSDEDDGGSEITGFIERDKAK 8742
Qy 91 ----- 90
Db 8743 MHTWRQPIETERSKCDITGLLEGOEYKFRVIAKNKFGGPPPEIGPILAVDPLGPTTSPE 8802
Qy 91 ----- 90
Db 8803 RLTYTERTKSTITLDWKEPRSGSPITQGYIIEKRRHDKDPDFRVNKRKLCPTTSELVNL 8862

QY 91 ----- 90
Db 8863 DEHQYEFVKAVNEIGESBPSPLNVVIQDDEVPTIKLRLSVRGDTIKVKACEPVHIP 8922
QY 91 ----- 90
Db 8923 ADVTGLPMPKIEMSKNETVIEKPTDALQITKEEVSSEAKTELSIPKAVREDKGTYTVA 8982
QY 91 ----- 90
Db 8983 SNRLGSVFRNVHVEYDRPSPRNLAVTDIKAESCYLTDAPLDNGGSEITHYVIDKRDA 9042
QY 91 ----- 90
Db 9043 SRKKAWEVNTAVAKRYGIWKLPNGQYEFVRVAVNKYGISDECKSDKVVIQDPYRLP 9102
QY 91 ----- 90
Db 9103 GPPGPKVLARTKGSMLVSWTPPLDNGGSPITGYWLEKREEGSPYWSRAPITKYGLK 9162
QY 91 ----- 90
Db 9163 GVEENVPRLLGVKYQFRAMAINAAGIGPPSESDPEVAGDPFPPGPPSCPVEVKDKTKS 9222
QY 91 ----- 90
Db 9223 SISLGWKKPAKGGSPIKGYIVEMQEEGTTDWKRVNPDKLITTCCECVWPNLKELRKYRF 9282
QY 91 ----- 90
Db 9283 RVKAVNEAGESEPSDTTGEIPATDIOBEVEFIDIGAODCLVCKAGSQIRIPAVIKGRPT 9342
QY 91 ----- 90
Db 9343 PKSSWEFDGKAKKAMKDCVHDIPEDAQLETAENSSVIIPECKRSHTGKYSITAKNKAGQ 9402
QY 91 ----- 90
Db 9403 KTANCRVKVMDVPGPKDLKVSIDITRGSRLSWKMPDDGGDRIGYVIEKRTIDGRAWT 9462
QY 91 ----- 95
Db 9463 KVNPDGSGTTFVVDLLSEQQYFFRVRARENRFIGIGPVETIQTARTDPIYPPDPPIKLK 9522
QY 96 ----- 95
Db 9523 IGLITKNTVHLSWPPKNDDGSPVTHYIVECLAWDPTGTKEAWRQCNKRDRVEELQFTVE 9582
QY 96 ----- 95
Db 9583 DLVEGGEYFRKAVNAAGVSKPSATVPCDCORPDPSPSIDLKEEFMEVEEGTNVIVAK 9642
QY 96 ----- 95
Db 9643 IKGVPFPTLTFWKAPPKPONKEPVLVDTHVNLVVDVDTCTLVIPQSRSSDTGLYTITAV 9702
QY 96 ----- 95
Db 9703 NNLTASKEMRLNVLGRPPVGPPIKFEVSADQMTLSWPPPKDDGSGKITNYVIEKREA 9762
QY 96 ----- 95
Db 9763 NRKTWVHVSSEPKCTYTIPLKLEGEVYFRIMAQNKYIGICEPLDSEPTARNLFSVPGA 9822
QY 96 ----- 95
Db 9823 PDKPTVSSVTRNSMTVNWEEPEYDGGSPVTGYWLEMKDOTTSKRWKVRNRPDIKAMTLGVS 9882
QY 96 ----- 95
Db 9883 YKVTGLIEGSDYQFRVAINAAGVPASLPSPDAPARDPIAPGPPFPKVTDTWKSSADL 9942
QY 96 ----- 95

Db 9943 EWSPLKDGGSKVTGYIVEYKEGKEBEWKGKDEVRGRTKLVTGLKEGAFYKFRVSAVN 10002
QY 96 ----- 95
Db 10003 IAGIGEPGEVTDVJEMKDRLVSPDLQLDASVDRIVVHAGGVIIRIIAYVSGKPPPTVTWN 10062
QY 96 ----- 95
Db 10063 MNERTLPQEAETIETTAISSMVKNCORSHOGVYSLAKNAEGERKKTIIIVDVLDPGPV 10122
QY 96 ----- 95
Db 10123 GTPELAHNLNESCKLTWFSPEDDGGSPITNYVIEKRESRRAMTPVTYTVTRQNAVQG 10182
QY 96 ----- 95
Db 10183 LIOGKAYFFRIAENSIGMGPFVETSEALVIREPITVPERPEDLEVKEVTNTVTLTWNP 10242
QY 96 ----- 95
Db 10243 PKYDGGSEIINYVLESRLIGTEKFHKYTNNDNLLSRKYTVKLGKEDTYEYRVSAVNIQG 10302
QY 96 ----- 95
Db 10303 GKPSFCTKPTCKDELAPPTLHLDFRDKLTIRVGEAFALTGRYSGKPKPKVSWFKDEADV 10362
QY 96 ----- 98
Db 10363 LEDDRTHIKTTPATLALAKIKAKRSDSGKYCVVVENSTGSRKGCQVNVVDRPDPGPV 10422
QY 99 ----- 98
Db 10423 SFDEVTKDYVVISWKPPDLDDGSGKITNYIIIEKVEGKDVWMPVTSASAKTCKVSKLLEG 10482
QY 99 ----- 98
Db 10483 KDYIFRIHAENLYGISDPLVSDSMKAKDRFRVPDAPDQPIVTEVTKDSALVTWNRKPHDGG 10542
QY 99 ----- 98
Db 10543 KPITNYILEKRETMKRWARTKDIPIHPYTKFRVPDLLEGQYEFVRSAENEIGIDPSP 10602
QY 99 ----- 98
Db 10603 PSKPVFAKDPTAKSPPVNPEAIDTTCNSVDLTWQPPRHDGSKILGYIVEYQKVGDSEW 10662
QY 99 ----- 104
Db 10663 RRANHTPESCPETKYKVTGLRDGQTKYFRVLAVNAAGESDPAHVPEPVLVKDRLEPPELI 10722
QY 105 ----- 104
Db 10723 LDANMARREQHIKVGDTLRLSAIIRKGVPPKVTWKEDRDAPTKARIDVTPVGSKLEIRNA 10782
QY 105 ----- 107
Db 10783 AHEDGGIYSLIVENPAGSKTVSVKVLVDKCPDRDLEVSEIRKDSCLYTWKEPLDDGGS 10842
QY 108 ----- 107
Db 10843 VITNVVVERRDVAQAQWSPLSATSKKSHFAKHLNEGNYQLFRVAANQYGRGPFVETPK 10902
QY 108 ----- 113
Db 10903 PIKALDPLHPPGPKDLHHVDVKTEVSLVWNNKPDORGSGPITGYLVEYQEGTDWIKF 10962
QY 114 ----- 113
Db 10963 KVTNLECVTGLQOGKTYFRVRAENIVGLDPTTPIECQEKLVPPSVELDKLIEG 11022
QY 114 ----- 113

Db 11023 LVVAGTTRFFPAIRGVPPVPTAKWTTDGGSEIKTDEHYTETDNFSSVLTIKNCLRRDTG 11082
QY 114 ----- 113
Db 11083 EYQITVNSAAGSKTAVHLTVLDVPGPTGPGPINILDVTEHMTISWQPKDDGGSPVINY 11142
QY 114 ----- 113
Db 11143 IVEKQDTRKDTWGVSSGSKTKLIPHLOKGEYVFRVAENKIGVPPLDSTPTVAKH 11202
QY 114 ----- 113
Db 11203 KFSPPSPGKPVVTDITENAAATVSWTLPKSDGSGPITGYIMERREVTGKWRVNKTPIAD 11262
QY 114 ----- 113
Db 11263 LKFRVTGLYEGNTYEFVRFAENLAGLSKSPSSDPIKACRPKPGPPINPKLKDKSRET 11322
QY 114 ----- 113
Db 11323 ADLWTKPLSDGSGPILGYVEQCPGTAQWNRINKDELIRQCAFVPGGLIEGNEYRPR 11382
QY 114 ----- 113
Db 11383 KAANIVGEGEPRELAESVIAKDILHPPEVELDVTCTRDVITVRVGTIRILARVKGREPED 11442
QY 114 ----- 113
Db 11443 ITWKEGVLVREKRVLDLIQDLPRVELOIKEAVRADHGKYYIISAKNSSCHAQSAIVNVL 11502
QY 114 ----- 113
Db 11503 DRPGPQNLKVTNVTKENCTISWENPLDNGSGEITNFIVEYRKPQKGWSIVASDVTKRL 11562
QY 114 ----- 113
Db 11563 IKANLLANNEYFRVCAENKVGVGPTIETKTPILAINPIDRGPENLHIADKGTFFYL 11622
QY 114 ----- 113
Db 11623 KWRRPDYDGGSPNLSYHVBERLLKSGDDWERVHKSGIKETHYMWDRVCENQIYEFVRQTKN 11682
QY 114 ----- 113
Db 11683 EGESDWWKTEEVVVKEDLQKPVLDLKLGSVLTVRAGDTIRLEAGVRGKPFPEVAWTKDK 11742
QY 114 ----- 113
Db 11743 DATDLTRSPRVKIDTRADSSKESLTAKRSDGGKYVVVTTATNTAGSFVAYATVNVLDKPGP 11802
QY 114 ----- 113
Db 11803 VRNLKIVDVSSDRCTVCWDPPEDDGGCEIQNYILEKCEKTRMWSTYSATVLTGTTVTR 11862
QY 114 ----- 113
Db 11863 LIEGNEYIFRVAENKIGTPTESKPVIAKTYDKPGRPDPPEVTKVSKEMTVVWNP 11922
QY 114 ----- 113
Db 11923 EYDGGKSITGYFLEKKEKSTRWVPVNKSAIPERRMKVQNLPLPDHEYQFRVKAENEIGIG 11982
QY 114 ----- 113
Db 11983 EPSLPSRPVAKDPIEPGPTNFRVDTKHSITLWGKPVYDGGAPILIGYVWEMRPI 12042
QY 114 ----- 113
Db 12043 ADASPEGMKRCNAAAQVLVRKEFTVTSLENOEYFRVCAQNOVGIGRPAELKAIPKE 12102
QY 114 ----- 113
Db 12103 ILEPPEIDLDSMRKLVIVRACCPILFAIVRGPAPKVTKVRKVGIDNVVRKGGQVLDVDT 12162

QY 114 ----- 113
Db 12163 MAFLVIPNSTRDDSKYSLTLVNPAGEKAVFVNVRLDTPGPPVSLKVSVDVTKTTSCHVSW 12222
QY 114 ----- 113
Db 12223 APPENDGGSQVTHYIVEKREADRKTWSTVTPEVKKTSFHVNTLVPGNEYFFRVTAVNEVG 12282
QY 114 ----- 113
Db 12283 PGVPTDVPKPVILASDPLSEDPDRKLEVTMTKNSATLAWPLRLDGGAKIDGYITSYRE 12342
QY 114 ----- 113
Db 12343 ESOPADRWTEYSVVVKDLSLVVTGLKEGKYYKFRVAARNAVGVSLPREAGVYEAKBQLLP 12402
QY 114 ----- 113
Db 12403 PKILMEQITIKAGKKLRIEAHVYKGPHTCKKKGEDEVVTSSEHLAVHKADSSSLITIK 12462
QY 114 ----- 113
Db 12463 DVTRKDSGYSLTAENSSGTDTKIKVVMVMDAPPPPPFDISDADACSLSHIPLBD 12522
QY 114 ----- 113
Db 12523 GGSNITNIVEKCDVSRGDMVTALASVTKTSRVGKLIPGOEYIFRVAENRFGISEPLT 12582
QY 114 ----- 113
Db 12583 SPKMQVAFPGFVGPSEPKNARVTKNKDCIFVANDRPDSGGSPILIGYLIERKERNLSLWV 12642
QY 114 ----- 113
Db 12643 KANDTLVRSTEYPCAGLVEGLEYSPIRIALNAGSSPPSKPEYVYVTAARMVDPPOKPEVI 12702
QY 114 ----- 113
Db 12703 DVTKSTVSLINARPKHDGSKIIGYFVEACKLPDGKWRCNTAPHOIQEEYATATGLEBK 12762
QY 114 ----- 113
Db 12763 AOYQFRAIARTAVNISPPSEPSDPTVILAENVPRIIDLSVAMKSLTVKAGTNVCLDATV 12822
QY 114 ----- 113
Db 12823 FGKPMPTVSWKKDGTLLKPAEGIKMAMORNLCITLLEFSVNRKDCDYTTITAENSSGSKSA 12882
QY 114 ----- 113
Db 12883 TIKLKVLDKPGPASVKINKMYSDRAMLSWEPPELGGSEITNYIVDKRETSRPNWAQVS 12942
QY 114 ----- 113
Db 12943 ATVPIITSGVEKLIBEGHEYQFRICAENKYGVGDPVFTEPAIAKNPYDPPGRCDPPVISNI 13002
QY 114 ----- 113
Db 13003 TKDHMTVSWKPPADGGSPITGYLLEKRETOAVNWKVNRKPIERTLKATGLQEGTEYE 13062
QY 114 ----- 113
Db 13063 FRVTAINKAGPKPSDASKAAVARDPOYPAPPAPFKVYDTRSSVSLSWGKPAYDGGSP 13122
QY 114 ----- 113
Db 13123 IIGYLVEYKRADSNWVRNCLPNQLQKTRFEVTGLMEDTOYQFRVYAVNKGIGYSDPSDVP 13182
QY 114 ----- 113
Db 13183 DKHYPKDILIPPEGELDADLRKTLILRAGVTMRILYVPVKGRPPPKITWSKPNVNLDRDIG 13242

QY 114 ----- 113
Db 13243 LDIKSTDFTLRCENVNKYDAGKYLITLNSCGKEYTIIVKVLDTGPPVNVVTKES 13302
QY 114 ----- 113
Db 13303 KDSAYVTWEPPIIDGGSPIINVVQKRDAERKSWSTVTTCCKTSFRVANLECKSYFFR 13362
QY 114 ----- 113
Db 13363 VFAENEYIGDPGETRDAVKASQTPGPVVDLKVRSVKSSCSIGWKPHSDGGSRIIGYV 13422
QY 114 ----- 118
NGKEY-----
Db 13423 VDFLTEENKQORVMKSLSLQYSAKDLTEGKEYTFRVSAENENGCTPSEITVVARDDDVA 13482
QY 119 ----- 118
Db 13483 PDLDLGLDLYLAKENSFRLKIPKGPAPSVSWKKGEDPLATDTRVSESSAVNTT 13542
QY 119 ----- 118
Db 13543 LIVYDCQSDAGKYTITLKNVAGTKEGTISIKVVGKPGIPTGPIKPFDEVTAEAMTLKWP 13602
QY 119 ----- 118
Db 13603 PKDDGSEITNYILEKRDSVNNKWTCASAVQKTTFRVTRLHEGMEYTFRVAENKYGVG 13662
QY 119 ----- 118
Db 13663 EGLSEPIVARHPFDPDAPPPNIVDVHRDSVSLTWTDPKKTGGSPITGYHLEFKERNS 13722
QY 119 ----- 118
Db 13723 LLWKRANKTPIMRDFKVTGLTEGLEYEFVRVMAINLAGVGKPSLPSEPVVALDIDPPGK 13782
QY 119 ----- 118
Db 13783 PEVINITRNSVTLIWTEPKYDGGHKLTYIVEKRDLPSKSMKANHVNPBCAFTVTDLV 13842
QY 119 ----- 118
Db 13843 EGGKYEFRIRAKTAGAISAPSESTETIICKDEYEAPTIVLDPTIKDGLTIKAGDTIVLN 13902
QY 119 ----- 118
Db 13903 AISILGKPLPKSSWSKAGDIRPSDITQITSTPTSSMLTIKYATRKDAGEYTIATNPF 13962
QY 119 ----- 118
Db 13963 TKVEHVKTVDVPPGPPVEISNVAEKATLTWTPLEDGGSPIKSVILEKRETSRLW 14022
QY 119 ----- 118
Db 14023 TVVSEDIQSCRHVATKLIQGNEXIFRVSANHYKGEPVQSEPVKMVDRFGPPGPEKPE 14082
QY 119 ----- 124
KCKVSN-----
Db 14083 VSNVTKNATVSWKRPVDDGSGSEITGYHVERREKSLRWVRAIKTPVSDLRCKVTGLQEG 14142
QY 125 ----- 124
Db 14143 STYEFVSAENRAGIPGPPSEASDVLKMDAAYPPGPPSNPHVTDTKKSASLANGKPHYD 14202
QY 125 ----- 124
Db 14203 GGLEITGVVHEQKVGEAWIKDTGTALRITQFVVPDQLTKENYFRISAINDAGVGE 14262
QY 125 ----- 130
KALPAP-----
Db 14263 AVIPDVELIVERMAPDELDLRLTLVVRAGLSIRIFVPIKGRPAPEVWTKNINLKN 14322
QY 131 ----- 130

Db 14323 RANIENTESEFTLLIIPECNRYDTGKFVMTIENPAGKKSGFVNVRLDTPGPVLNLRPTDI 14382
QY 131 ----- 133
-IEK-----
Db 14383 TKDSVTLHWDPLIDGGSRTNYIVEKREATRKSYSSTATTKCHKCTYKVTGLSGEGCEYFF 14442
QY 134 ----- 133
-:|-----
Db 14443 RVMAENEYIGEPTEPETTEPVKASEAPSPDLSINIMDITKSTVSLAWPKPKHGGSKITGY 14502
QY 134 ----- 133
Db 14503 VIEAQRKGDOWHTHTTVKLECVVRNLTEGEYTFQYMAVNSAGRSAPRESRPVIVREQ 14562
QY 134 ----- 133
Db 14563 TMLPELDLRGIYQKLVIAKAGDNIKVEIPVLGRPKPTVWKKGDOIILKQTORVNFETTAT 14622
QY 134 ----- 133
Db 14623 STILNINECVSDSGPYPLTARNIVGEVDVITIQVHDI PGPTGPIKFEVSDSDEVTF 14682
QY 134 ----- 133
Db 14683 WDPPEDDGVPISNVWVEMROTDSTTWVELATTVIRTTTKATRLTTGLEYOFRVKAQNR 14742
QY 134 ----- 133
-TISKA-----
-:|:|-----
Db 14803 RNLGILWQTVSKALVPGNIFKSSGLTDGIAYEFVIAENMAGSKSPKSEPMALDIDP 14862
QY 139 ----- 138
Db 14863 PGKVPNLNTRHTVTLKWAKEPYTGGFKITSIVEKRDLPNGRWLKNFNLNENFTVS 14922
QY 139 ----- 138
Db 14923 GLTEADAAEFVRVIAKNAAGAI SPPSEPSDAITCRDDVEAPKIKVDVKFKDVTILKAGEAF 14982
QY 139 ----- 138
Db 14983 PLEADVSGRPPPTMEWSKDGKELEGTAKLEIKIADFSNLNKNKSTRDSDGAYTLTATNP 15042
QY 139 ----- 138
Db 15043 GGFAKHIFNVKLDVRPGPEGLAVTEVTSEKCVLSWPPPLDDGGAKIDHVIVQKRETSR 15102
QY 139 ----- 138
Db 15103 LAWTVASEVOVTKLVTKLLKGNEYIFRVMVNNKYGGEPLSEPFVLAVNPYPGPPDPK 15162
QY 139 ----- 138
Db 15163 NPEVTTITKDSMVVCGWHPDSDGSGSEIINYIVERREDKAGQRIKCNKKTLLDRLRYKVSGL 15222
QY 139 ----- 138
Db 15223 TEGHEYEFRI MAENAGISAPSTSPFYKACDVTVPKPGPNPVRDLTSSRSISIAWNP 15282
QY 139 ----- 138
Db 15283 IYDGGSEITGYMVEIALPEDEWQIVTPPAGLKATSYTITGLTENQEKYIRIYAMNSGL 15342
QY 139 ----- 138
Db 15343 GEPALVPCTPKAEDRMPLPELADLRKVVTIRACCTLRFLVPIKGRPAPEVKNWARDHG 15402
QY 139 ----- 138

Db 15403 ESLDKASIESTSSYTLILVGNVNRFDGKGYILTVENSNGSKSAFVNVVRVLDTPGPPQDLK 15462
QY 139 ----- 138
Db 15463 VKEVTKTSVTLWDPPLLDGGSKIKNYIIVEKRESTRKAYSTVATNCHKTSWKVDQLQEGC 15522
QY 139 ----- 138
Db 15523 SYFRVLAENEYIGLPAETAESVKASERPLPGKITLMDVTRNSVLSWEKPEHDGGR 15582
QY 139 ----- 138
Db 15583 ILGYIVEMOTGSKDWATCATVKVTEATITGLIQGEYSFRSAQNEKGISDPRQLSVPV 15642
QY 139 ----- 138
Db 15643 IAKDLVIPAPAKLLFNFTFVLAGEDLKVDPFPIGRPTPAVTHHKDNVPLKQTRVNAEST 15702
QY 139 ----- 138
Db 15703 ENNSLLTIKDACREDGVHVKLTNSAGEAIELNVLDKPGPTGPKVMDVETADSIT 15762
QY 139 ----- 138
Db 15763 LSWGPPKYDGGSSINNYIIVEKRDSTTTWQIVSATVARTTIKACRLKTCGEYQFRIAEN 15822
QY 139 ----- 138
Db 15823 RYKSTYLNSEPIVAQYPKVPCPGTPVVTLSRDSMEVQNEPISDGGSRVIGYHLER 15882
QY 139 ----- 138
Db 15883 KERNSILWKLKNTPIPOTKFKTTGLEGEVEFRVSAENIVGIGKPSKVCYVARDPC 15942
QY 139 ----- 138
Db 15943 DPPGRPEALIVTRNSVTLQWKKPTVDGSKITGYIIVEKKELPEGRMMKASFNTIIDTHPE 16002
QY 139 ----- 138
Db 16003 VTGLVEDHRYEFRVIARNAAGVFSESESTGALTARDEVDPPIRISMDPKYKDTIVVHAGE 16062
QY 139 ----- 138
Db 16063 SFKVDADIYKPIPTIQWIKGQDLSNTARLEIKSTDFATSLSKDAVRVDSGNVILKAK 16122
QY 139 ----- 138
Db 16123 NVAGERSVTNVKVLDRPGPPGPPVVISGVTAEKCTLAWKPLQDGGSDIINIYIVERRET 16182
QY 139 ----- 138
Db 16183 SRLVWTVVDANVQTLSCVKYTKLLEGNEYFRIMAVNKYGVGEPLSEPVVAKNPFVPPDA 16242
QY 139 ----- 138
Db 16243 PKAPEVTTVTKDSMIVWVERPASDGGSEILGYVLEKRDKEGIRWTRCHKRLIGELRLRYT 16302
QY 139 ----- 138
Db 16303 GLIENHDEFRVSAENAGLSEPPSPSAVQKACDPIYKPGPPNPKVIDITRSSVFLSWS 16362
QY 139 ----- 138
Db 16363 KPIYDGGCEIOGYIIVEKCDVSGVETMCTPTGINKNTNIEVEKLEKHEYNEFRICAINKA 16422
QY 139 ----- 138
Db 16423 GYGEHADVPPIVEKLEAPDIDLDELRLKIINIRAGGSLRLFVPKIGRPTPEVKWGV 16482
QY 139 ----- 138
Db 16483 DGEIRDAAIIDVTSFTSLVLDNVNRYDSGKYTLTLENSSGTKSAFVTVRVLDTPSPPVN 16542

QY 139 ----- 138
Db 16543 LKVEITKDSVSTWEPPLLDGGSKIKNYIIVEKREATRKSAAVVTNCHNSHKIDLOQE 16602
QY 139 ----- 138
Db 16603 GCSYFRVTAENEYIGLPAQTAADPIKVAEVPQPPGKITVDDVTRNSVLSWTKPEHDGG 16662
QY 139 ----- 138
Db 16663 SKIIQIVEMOAKHSEKSECARVKSLOAVITNLTOGEEYLFRRVAVNEKGRSDPRSLAV 16722
QY 139 ----- 138
Db 16723 PIVAKDLVIEPDVKPAFSSYSVQGDLEIIEVPIISGRPKPTTWTGDLPLKQOTTRINVT 16782
QY 139 ----- 138
Db 16783 DSLDLTTLTSIKETHKDDGGGYITVANVYGQKTASIEIVTLDKPPPKGVKFDVSAES 16842
QY 139 ----- 138
Db 16843 ITLSWNPPLYTGCCQITNYIVOKRDTTWTWVVSATVARTTLKVTKLATGTEYQFRI 16902
QY 139 ----- 138
Db 16903 ENRYGOSFALES DPIVAQYPYKEPGPGTFFATAISKDSWVQWHEPVPVNGGSPVIGYHL 16962
QY 139 ----- 138
Db 16963 ERKERNILWTKVNTKTIHDTQFKAQNLBEGIEYEFVRVYAENIVGVKASKNSECYVARD 17022
QY 139 ----- 138
Db 17023 PCDPPTPEPIMVKRNEITLQWTKPVYDGGSMITGYIVEKRDLPGRMMKASFNTVNETQ 17082
QY 139 ----- 138
Db 17083 FTVSGLTEDQRYEFRVIAKNAAGAIKSPSDSTGPIAKDEVELPRISMDPKFRDITVUNA 17142
QY 139 ----- 138
Db 17143 GETFRLEADVHGKPLPTIEWLRGDKIEESARCEIKNTDFKALLIVKDAIRIDGGYILR 17202
QY 139 ----- 138
Db 17203 ASNVAGSKSPVNVKVLDRPGPPGPPVQVGTSEKCSLTWSPPLQDGGSDISHYVVEKR 17262
QY 139 ----- 138
Db 17263 ETSRLANTVVASEVVTNSLKVTKLLEGNEYVFRIMAVNKYGVGEPLSEAPVLMKNPFVLP 17322
QY 139 ----- 138
Db 17323 GPPKSLEVTNIAKDSMTVCWNRPDSDGGSEIIGYIVEKRDGIRWIKCKRKRITDLRLR 17382
QY 139 ----- 138
Db 17383 VTGLTEDHEYEFVRVSAENAGVGEPSATVYKACDPVKPGPPTNAHIVDTTKNSITLA 17442
QY 139 ----- 138
Db 17443 WCKPIYDGGSEILGYVVEICKADEEWOIVTPTQGLRVTRFEISKTEHQEKIRVCALN 17502
QY 139 ----- 138
Db 17503 KVGLEATSVPGTVKPEKLEAPDLDBSELKGIWRAGGSARHIPFKGRPTPEITWS 17562
QY 139 ----- 138
Db 17563 REGEFTDKVIEKGVNTQLSIDNCDNRDAGKYILKLENSSGSKSAFVTVKVLDTPGPP 17622

QY 139 ----- 138
Db 17623 QNLAVKEVRKDSAFLEWPEPIIDGGAKNVVIDKRESTRKAYANVSSKCKTSFKVENL 17682
QY 139 ----- 138
Db 17683 TEGAIYFRYMAENEFVGVPVETDVAKAAEPSPGKVTLLTDSOTSASLWKEPEHD 17742
QY 139 ----- 138
Db 17743 GGSRLGVVYVEMQPKGTEKHSIIVAESKVCNAVVTGLSSGOEYQFRVKAYNEKKSDBRVL 17802
QY 139 ----- 138
Db 17803 GVPVIAKDLFIQPSLKLPPNTYSIQAGEDLKIEIPVIGRPPRPNLSMWKDGLEPLKQTRVN 17862
QY 139 ----- 138
Db 17863 VEETATSTVLHIKEGNKDDFGKTYTATNSAGTATENLSVIVLEKPGPPVGVRFDEVSA 17922
QY 139 ----- 138
Db 17923 DFVVISWEPAYTGGCOISNIYVEKRDTTTTWHMVSATVARTIKITKLKTGETYQFRI 17982
QY 139 ----- 138
Db 17983 FAENRYKGSAPLDSKAVIQVPFKEPGPGTPTFTSISKQMLVQWHEPVNDGGTKIIGY 18042
QY 139 ----- 138
Db 18043 HLEQKEKNSILWKLAKNTPIDQTKFKTTGLDEGLEEFYKVAENIVIGRPSKVSECFVA 18102
QY 139 ----- 138
Db 18103 RDCDPPRPEAIIVTRNNVTLKWKPAYDGGSKITGYIVEKKDLDPGRMKASFNVLE 18162
QY 139 ----- 138
Db 18163 TEFTVSLVEDQRYEPRVIAANAAGNSEPSDSSGALTARDEIDAPNASLDPKYKDVIVV 18222
QY 139 ----- 138
Db 18223 HAGETFVLEADIRGKPIPDVVWSKDGKELEETAARMEIKSTIOKTLVLVVKDCINTDGGQY 18282
QY 139 ----- 138
Db 18283 ILKLSNVGGTKSIPITVKVLDLDRPGPEGLKVTVTAEKCYLAWNPPLODGGANISHYII 18342
QY 139 ----- 138
Db 18343 EKRETSRLSWTQVSTEYOALNYKVTKLLPGNEYIFRVMVKNYKIGEPLESGPVTACNPY 18402
QY 139 ----- 138
Db 18403 KPPGPPPTPEVSAITKDSMVVTWARPVDDGGTEIEGYILEKRDKEGVRTKCNKKTLDL 18462
QY 139 ----- 148
Db 18463 RLRTVGLTEGHSYEFYVAENAAGVGEPSVFIYACDALYPGPPSNPKVTDTSRSSV 18522
QY 149 ----- 148
Db 18523 SLAWSKPIYDGGAPKGVYVVEKAAADEWTTCTPTPTLOGKQFTVTKLKENTENPRIC 18582
QY 149 ----- 151
Db 18583 AINSEGVGEATLPUGSVVAQERIEPPELADLRKVVVLRASATLRLFTVIRGRPEPEV 18642
QY 152 ----- 151
Db 18643 KWEKAEGILTDRAGIEVTSSTMLVIDNVTRFDSGRYNLTLENNSGSKTAFVNVRLDSP 18702
QY 152 ----- 151

Db 18703 SAPVNLITREVKKDSVTLWEPPILIDGGAKITNYIVEKRETRKAYATITNCTKTFR 18762
QY 152 ----- 151
Db 18763 ENLOBCSSYFRVLASNEYIGLPAETTEPVKVSEPLPPGRVTLVDVTRNTATIKWEKP 18822
QY 152 ----- 151
Db 18823 ESDGSKIITGVVYVEMQTKSEKWSCTQVKTLEATISGLTAGEEYFVRAAVNEKGRSDP 18882
QY 152 ----- 151
Db 18883 RQLGVPVIARDIEIKPSVELPFHTFNVKAREQLKIDVPFKGRFOATVNRKDGOTLKETT 18942
QY 152 ----- 151
Db 18943 RVNVSSKTVTSLSIKEASKEDVGYELCVSNSAGSTVPTIIVLDRPDPGPIRIDEV 19002
QY 152 ----- 151
Db 19003 SCDSTITSNWPPEDGCGCOISNIYVEKKTSTTWHIVSOAVARTSIKIVRLTGTSEYQF 19062
QY 152 ----- 151
Db 19063 RVCAENRYGKSSYSESAVVAEYFPSPGPGTPKVVHATKSTMLVTWQVPVNDGGSRV 19122
QY 152 ----- 151
Db 19123 GYHLEYKERSIILWSKANKILIAQTOMKVSGLDEGLMEYRYVAENIAGIKCSKCEPV 19182
QY 152 PSRD----- 155
|:|
Db 19183 PARDCPPCQPEVTNITRKSLSKWSKPHYDGGAKITGYIVERRELPGRWLKCNTNI 19242
QY 156 ----- 155
Db 19243 QETYFEVTELTEDQRYEFVRFARNAADSVSEPSSTGPIIVKDDVEPRVMDVKFRDVI 19302
QY 156 ----- 155
Db 19303 VVKAGEVLKINADIAGRLPVISWAKDGIIEERARTEIISTDNHLLTVKDCIRRTGQ 19362
QY 156 ----- 155
Db 19363 YVLTKNVAGTRSVAVNCKVLDKPGPPAGPLEINGLTAEKCSLSWGRPOEDGGADIDYI 19422
QY 156 ----- 155
Db 19423 VEKRETSHLAWTICEGLOMTCKVTLLKAGNEYIFRVTGVNKGVGEPLESAIKALDP 19482
QY 156 ----- 155
Db 19483 FTVPSPPTSLEITSVTKESMTLWRSRDESDGSEISGYIITERREKNSLRVVRNKKPVYD 19542
QY 156 ----- 155
Db 19543 LRVKSTCLREGCEYEYRYVAENAGLSLPSETSPHIRAEDPVFLPSPPKPIVDSCKTT 19602
QY 156 ----- 155
Db 19603 ITIAWVKPLFDGGAPITGYTYEYKKSDTDWKTSIQSLRGTEYTSISGLTTGAEYVFRVKS 19662
QY 156 ----- 155
Db 19663 VNVKASDPSSDPQIAKEREEREPFLDIDSEMRKTLIVKAGASFTVTPFRGRPVNVL 19722
QY 156 ----- 155
Db 19723 WSKPDTDLTRAYVDITDSRTSLTIENANRNDCKYTLTIQNVLSAASLTLVVKVLDTPG 19782
QY 156 ----- 155

Db 19783 PPTNITVQDVTKEAVLSWDVPENDGAPVKVNYHIEKREASKAAVSVTNNCNRLSYKYT 19842
QY 156 ----- 155
Db 19843 NLQEGAIYPRVSGENEFVGIPAETKEGVKITEKSPPEKLGVTISKXDSVSLTWLKEP 19902
QY 156 ----- 155
Db 19903 HDGSRIVHYVVALEKQKNWVKCAVAKSTHHVSGLRSENSEYFFRFAENQAGLS DPR 19962
QY 156 ----- 155
Db 19963 ELLLPVLIKEQLEPPEIDMKNFPSHTVYVRAGSNLKVDPISGKPLPKVTLSDRGVPLKA 20022
QY 156 ----- 155
Db 20023 TMRNTEITAENLTINLKESVTADAGRYEITAANSSTTKAFINIVVLD RPPGTPPV 20082
QY 156 ----- 155
Db 20083 SDITEESVTLKWEPPKYDGGSQVTNYILLKRETSTAVMTEVSATVARTMMKVMKLTGEE 20142
QY 156 ----- 155
Db 20143 YQPRIKAENRFGISDHDSACVTVKLPYTPGPPSTPWNTVNTRESITVGNHEPVSNGS 20202
QY 156 ----- 155
Db 20203 AVVGYLEMKDRNSILWQANKLVIRTHFKVTITISAGLIYEFYVAENAGVGRPSHPS 20262
QY 156 -----ELTNQVSL----- 164
Db 20263 EPVLADACEPPRNRVITDISKNSVLSWQPAFDGSGKITGYIVERDLDPGRWTKASF 20322
QY 165 -----TCL----- 167
Db 20323 TNVTETQFIISGLTONSQYEFYRFAVNAVGSISNPSEVVGPTICDSYGGPVIDLPLEYT 20382
QY 168 ----- 167
Db 20383 EVVKYRAGTSVKLRAGISGKPAPTIEWYKDKELQTNALVCVENTDLASILIKDADRLN 20442
QY 168 ----- 167
Db 20443 SGCYELKLRNAMSASATIRVQILDKPPGPGPIEFKVTVAEKITLLWRPPADGGAKIT 20502
QY 168 ----- 167
Db 20503 HYIVEKRETSRVVMSVSHLEECITTTKIIKGNEYIFRVRVANKYIGEPLESDSVVA 20562
QY 168 ----- 167
Db 20563 KNAFTPGPGPIPEVTKITKNSMTVWVRPIADGSDISGYFLEKRDKSLGFWKVLKET 20622
QY 168 ----- 167
Db 20623 IRDTROKVTGLTENSIOYRVCAVNAAGQPPSEPFYKAADPIDPPGPAKIRIADST 20682
QY 168 ----- 167
Db 20683 KSSITLGSKPVYDGGSAVTGYVVEIROGEEETVSTKGEVRTTEYVVSNLKPGVNY 20742
QY 168 ----- 167
Db 20743 FRVSAVNCAGQGEIEMNEPVQAKDILEAPEIDLVALRTSVIAKAGEDVQVLIPFKGRP 20802
QY 168 ----- 167
Db 20803 PPTVTRKDEKNGSDARYSIENTDSSLLTIPQVTRNDTGKYLITL IENGVPKPSVTS 20862
QY 168 ----- 167
Db 20863 VKVLDTPAACQKLVKHYVSRGTVLLWDPPLIDGSPINIVVIEKRDATKRTWSVVS HKC 20922

QY 168 ----- 167
Db 20923 SSTSFKLIDLSEKTPFFFRVLAENEIGIGPCETTEPVKAAEVPAPIRLSMKDKSTKTSV 20982
QY 168 ----- 167
Db 20983 ILSWTKPDFDGGSVITEYVVERKKGEOQWSHAGISKTCEIEVSQLEQSVLEFRVFAKN 21042
QY 168 ----- 167
Db 21043 EKGLSDPVTIGPITVKELIITPEVDLSOIPGAQVTVRIGHNVHLELPYKGPKPSISWLK 21102
QY 168 ----- 167
Db 21103 DGLPLKESEFVRSKTENKITLSIKNAKKEHGKYTVILDNAVCRITAVPITVITLGPSPK 21162
QY 168 ----- 167
Db 21163 PKGPIRFDEIKADSVILSDVPEDNGGGBEITCYSIEKRETSQTNWKMVCSSVARTTFKVP 21222
QY 168 ----- 167
Db 21223 NLVKDAEQYQFVRAENRYGVSOPLVSSIIIVAKHQFRIPOGPGKPYIYNVTS DGMSLTWD 21282
QY 168 -----VKGFY----- 172
Db 21283 PVDGSGSEVTGFHVEKKEKERNILWQKVNTSPISGREYRATGLVEGLDYQFVYAENSAGL 21342
QY 173 ---PSD----- 175
Db 21343 SSPSPDSKFTLAVSPVDPGTPDYIDVTRETITLKNPPLRGGSKIVGYSIEKROGNER 21402
QY 176 ----- 175
Db 21403 WVRCNFTDVSECQYTVTGLSPGDRYEFRIARNAVGTISPPSSGSIIMTRDENVPPIVE 21462
QY 176 ----- 175
Db 21463 FGPEYFDGLI KSGESLRILKALVQGRPVPRVTFWFKDGVIEKRMNWEITDVLGTSLSFVR 21522
QY 176 ----- 175
Db 21523 DATRHRGVYTVAEAKNAGSAAKAEIKVKQDTPGKVGVPIRFTNITGEKMTLWMDAPLND 21582
QY 176 ----- 175
Db 21583 GCAPITHYIIEKRETSRLAWALIEDKCEAQSYTAIKLINGNEYQFRVSAVNNFGVGRPLD 21642
QY 176 ----- 175
Db 21643 SDPVVAQIOYTVPDAPGIPEPSNITGNSITLTWARPESDGGSEIOQYILERREKSKTRW 21702
QY 176 ----- 175
Db 21703 KVISKRPISERFKVTGLTEGNEYEFHVMAENAGVGPASGISRLIKCREPVNPPPTV 21762
QY 176 -----TAVEW----- 180
Db 21763 VKVTDTSKTTVSELSKVPFDGMBEIIIGYI IEMCKADLGDWHKVNAEACVKT RYTVTDLQ 21822
QY 181 ----- 180
Db 21823 AGEYKFRVSAINGAGKDSCEVTGTIKAVDRLTAPELIDIDANFKQTHVVRAGASIRLFI 21882
QY 181 -----ESN----- 183
Db 21883 AYQGRPTPAVMSKPDNSLRLADIHTTDSFSTLTVCNCRNDAGKYTLTVENN S GSKSI 21942
QY 184 ----- 183
Db 21943 TFTVKVLDTPGPPGPIPFKDVTRGSATLWMDAPLIDGGARITHHYVVEKREASRRSQVVIS 22002

QY 184 ----- 183
Db 22003 EKCTROIFKVNDLAEGVPIYFRVSANVEYGVGEPIEMPEPIVATEQAPPRRLDVDTSK 22062
QY 184 ----- 183
Db 22063 SSAVLAWLKPDDHGGSRITGYLLEMRQKGSDFWEAGHTKQLTFTVERLVEKTEYEFVRK 22122
QY 184 ----- 183
Db 22123 AKNDAGYSEPREAFSSVIIKEPOIEPTADLTGITNOLITCKAGSPFTIDVPSGRPAKV 22182
QY 184 ----- 183
Db 22183 TWKLEEMRLKETDRVSITTTKDTTLTVKDSMRGDSGRFLTLTENTAGVKTFSVTVVIG 22242
QY 184 ----- 186
QY 184 -----GQP-----
|:|
Db 22243 RGPVGTGPIEVSSVAESCVLWSGEPKGGTEITNYIVEKRESGTTAWOLVNSSVKRTQ 22302
QY 187 ----- 186
Db 22303 IKVTHLTKWYFSRVSSNRFVSKPLSAPIIAEHFPVPPSAPTRPEYHVHSANAMSI 22362
QY 187 ----- 193
QY 187 -----ENNYKTT-----
|:|
|:|
|:|
Db 22363 RWEPEYDGGSKIIGYWEKERTILWKENKVPCLCNYKVTGLVEGLEYOFRYALN 22422
QY 194 ----- 193
Db 22423 AAGVSKASEASRPIMAQNPVDPAGRPEDVTRSTVLSIWSAPADYDGGKVVGYIIERKP 22482
QY 194 ----- 193
Db 22483 VSEVGGRWLKCNYTIVSDNFFVTVALSEGDYEFVRVLAKNAAGVTSKSESTGPVTCRD 22542
QY 194 ---pp-----
|:|
|:|
Db 22543 EYAPPAELDARLHGLVTRAGSDLVDAAVGGKPEPKIITWKGDKELDLCEKVSLOQT 22602
QY 196 ----- 201
QY 196 -----VLDSG-----
|:|
|:|
|:|
Db 22603 GKRAVIFKCDRSDSGKTYLTAVKNASGTVKAVSVMKVLDSPGCPGKLTAVSRVTQEKCTL 22662
QY 202 ----- 201
Db 22663 ANSLPOEDGGAETHVIVERRETSRLNWVIVEGECPTLSYVVTRLIKNEYIFRVRVAVNK 22722
QY 202 -----SFFL-----
|:|
|:|
Db 22723 YGPVPESEPIVARNSTIPSPGPIEVEGTGKEHIIQWTKPESDGGNEISNVLVDKR 22782
QY 206 ----- 205
Db 22783 EKSLRWKNDYVYDTRKVTSLMEGCDYQFRVAVNAAGNSEPSEASNFISCREPS 22842
QY 206 ----- 205
Db 22843 YTPGPPAPRVDDTKHSISLAWTKPMYDGGTDIVGYVLEMQEKDQDWVRVHTNATIRN 22902
QY 206 ----- 205
Db 22903 TEFTVPDLKMGQYSFRVAAVNKGHSESAETIEPVERIEIPDLELADDLKKTVTIR 22962
QY 206 ----- 213
QY 206 -----YSKLTVDK-----
|:|
|:|
|:|
Db 22963 AGASRLWVSVGRPPPVITWSQKIDLASRAIIDTTESYSLLIVDKVNRDAGKYTIEA 23022
QY 214 ----- 213
Db 23023 ENQSGKKSATVLVKVYDTGCPGCVKVEVSRDSVTITWEIPTDGGAPVNNVIVEKREA 23082
QY 214 ----- 213

Db 23083 AMRAFKVTTKCSKTLYRISGLVEGTWYFRVLPENIYGIGPECETSDAIVLSEVPLVPA 23142
QY 214 ----- 213
Db 23143 KLEVVDVTKSTVTLAWEKPLYDGGSRITGYVLEACKRAGTERMMKVVTLPKPTVLEHTVTSL 23202
QY 214 ----- 213
Db 23203 NEGEQYLFRIRAONEKGVSEPRETVTAVTVDLRLVLTIDLSTMPQKT IHVPAGRPVELV 23262
QY 214 ----- 213
Db 23263 IPIAGRPPPAASWFFAGSKLRESERVIVETHTKVAKLTIRETTIROTGEYTLLELKNVTGT 23322
QY 214 ----- 213
Db 23323 TSETIKVIIIDKPPPTGPIKIDEIDATSITISWEPPELDGGAPLSGYVVVEQDAHRPGW 23382
QY 214 ----- 213
Db 23383 LPVSESVTRSTFKFTRLTEGNEVYFRVAATNRRFGIGSYLQSEVIECRSSIRIPGPETLQ 23442
QY 214 ----- 213
Db 23443 IPDVSRDGMTLTWYPPEDDDGSOVTGYIVERKEVRADRVVRVNVKVPVTMTRYRSTGLTEG 23502
QY 214 ----- 213
Db 23503 LEYEHVRVTAIARGSGKPSRSPKPIVAMDP IAPPGKPNQPRVDTTTRTSVLAWSPEDE 23562
QY 214 ----- 213
Db 23563 GSKVTGYLIEMQVQDHEWTKCNTPTTKIREYTLTHLPQGAERYRVLACNAGGPEPA 23622
QY 214 -----SRMQOQ-----
|:|
|:|
|:|
Db 23623 EYVGTVKVTEMLEYPDYELDERYOEGIFVRQGGVIRLTIPKGPFPCKWKTEGODISK 23682
QY 220 ----- 219
Db 23683 RAMIATSETHTELVIKEADRGDSGTVDLVLENKCGKKAIVIKRVITGSPNSPEGPLEYDD 23742
QY 220 ----- 219
Db 23743 IQVRSVRSWRPPADGGADILGYILERREVPKAAWTTIDSRVRCGTSLVVVKGLKENVEYH 23802
QY 220 ----- 219
Db 23803 FRVSAENQFGISKPLKSEEPVTEKTLNPPPEPPNPPELVDVTKSSVLSWSRPKDDGGS 23862
QY 220 ----- 219
Db 23863 RVTGYIERKETSTDKWRHNTQITTTMYTVTGLVPDAEYQFRIIAQNVDVGLSETSPAS 23922
QY 220 ----- 219
Db 23923 EPVVCCKDPDKPSQPGELEILSISKDSVTLOWEKPECDGKGEILGYWVEYRQSGDSANWK 23982
QY 220 ----- 219
Db 23983 SNKERIKDKQFTIGGLLEATEYEFVRFAENETGLSRPRRTAMSIKTTLTSGEAPGIRKEM 24042
QY 220 ----- 219
Db 24043 KDVTTKLGEAAQLSCQIVGRPLPDIKWRFKELIOSRKYKMSDGRTHLTVMTEEQED 24102
QY 220 ----- 219
Db 24103 EGVYTCIATNEVEVETSSKLLLOATPQPHPGYPLKEKYYGAVGSTLRLHVMYIGRPVPA 24162
QY 220 ----- 222
NVF-----
|:|

```
Db 24163 MTWFHQKLLQNSENITIENTEHTYHLVKNVQRKTHAGKYKVQLSNVFGTVDAILDVEI 24222
QY 223 ----- 222
Db 24223 QDKPKDPTGPIVIEALLKNSAVISKPPADGGSWITNVVVEKCEAKEGAEQWLVSSAIS 24282
QY 223 ----- 222
Db 24283 VTTCRIVNLTENAGYVFRVSAQNTFGISDPLEVSSWIIKSPFEKPGAPGKPTITAVTKD 24342
QY 223 SCSV----- 226
      || |
Db 24343 SCVWAMKPPASDGGAKIRNYYLEKREKKONKWISVTTEIRETFVSVKNLIEGLEYEFRV 24402
QY 227 ----- 226
Db 24403 KCENLGGESENSEISEPTIPKSDVPIQAPHKEELNRLNRYQSNATLVCKVTGHPKPIV 24462
QY 227 ----- 226
Db 24463 KWYROKEIIDGLKYRIOEFKGGVHQLIIASVTDDATVYQVRATNOGGSVSGTASLEV 24522
QY 227 ----- 226
Db 24523 EYPAKIHLPKLTLEGMGAVHALRGEVSIKIPFGKPDPIITWQKGQDLIDNNGHYQVIVT 24582
QY 227 ----- 226
Db 24583 RSFTSLVPNGVERKDAGFYVVCARNREGIDQKTVELDVADVPDPGRGVKVSVDVSRDSVN 24642
QY 227 ----- 226
Db 24643 LTWTEPASDGGSKITNYIVEKATTAERWLRVGOARETRYVINLFGKTSYQFRVIAENK 24702
QY 227 ----- 226
Db 24703 FGLSKPSEPTITTKEDKTRAMNYDEEDETREVSMTKASHSTKELYERYMIAEDLGR 24762
QY 227 ----- 230
      -MHEA-
      :||:
Db 24763 GEFGIVHRCVETSSKKTYMAFVKVKGTDQVLVKKESILNIAHRNHLHESFESMEE 24822
QY 231 -----LNHN----- 234
      ||:|
Db 24823 LVMIFISGLDIFERINTSAFELNEREIVSVHQVCEALQFLSHNIGHFDIRPENIY 24882
QY 235 ----- 234
Db 24883 QTRRSSTIKIIEFGQAROLKPGDNFRLLFTAPEYYAPEVHQHDVVSTATDMWSLGLTVYV 24942
QY 235 -----YT----- 236
      ||
Db 24943 LLSGINPFLAETNQOIENINNAEYTFDEAFKEISIEAMDFDRLLVKERSMTASEA 25002
QY 237 -----QKSLSLS- 243
      || :||:
Db 25003 LQHPWLKOKIERVSTKVRTLKHRRYHTLIKDLNMVMSAARISCGGAIRSQRGSVAK 25062
QY 244 ----- 243
Db 25063 VKVASIEIGPVSGQIMHVAEGEGGHVKYCKIENYDOSTQVTWYFGVQLENSEKYEITY 25122
QY 244 ----- 243
Db 25123 EDGVAILVVKDITKLDDGTYRCKVVNDYGEDSSYAELFVGKGVREYDYCYCRTMKIKRR 25182
QY 244 ----- 243
Db 25183 TDTMLLERPPEFTLPLYNKTYAYGENVRFGVITVHPHPTWYKSGQIKPGDNDKKY 25242
QY 244 ----- 243
Db 25243 TFSDKGLYQLTINSVTDDDAEYTVVARNKYGEDSCAKLTVTLHPPPTDSTLRPMFKR 25302
```

```
QY 244 ----- 243
Db 25303 LLANACQEGOSVCFEIRVSGIPPTLKWEXDQGPLSLGPNIEIIHEGLDYVALHIRDTL 25362
QY 244 ----- 243
Db 25363 PEDTGYRVTTATNTAGSTSCQAHLQVERLRYKQBEFKSKEGHERHVQKIDKTLRMAEIL 25422
QY 244 ----- 243
Db 25423 SGTESVPLTQVAKREALREAAVLRYKPAVSTKTVKGEFRLEIEBKKEERKLMPYDVPEPRK 25482
QY 244 -----PGK 246
      |||
Db 25483 YKQTTIEEDQRIKQFVPMDSMDMKYKKIRDOYEMPGK 25518
```

Search completed: April 21, 2003, 10:50:13
Job time : 99.1765 secs

PD 29-JUN-2000.
XX
PF 23-DEC-1999; 99WO-US30900.
XX
PR 23-DEC-1998; 98US-0113387.
XX
PA (REGE-) REGENERON PHARM INC.
XX
PI Davis SJ, Gale NW, Yancopoulos GD, Stahl N;
XX
PI WPI; 2000-442670/38.
DR N-PSDB; AAA51345.
DR
XX
PT Polynucleotide encoding a fusion polypeptide, useful for promoting
PT differential function and influencing phenotype, comprises two subunits
PT containing at least one copy of the receptor binding domain of a ligand
XX
XX Example 12; Fig 14A-E; 97pp; English.
PS
XX
CC Production of homogenous forms of clustered ligands is broadly applicable
CC to improve the affinity and/or increase the activity of a ligand as
CC compared to the native form of the ligand. Ephrin fusion proteins have
CC been constructed, which may be useful for treating neurological
CC disorders. The ephrin fusion proteins are preferably capable of binding
CC to Etk receptor and are especially Etk-6 antagonists. The fusion proteins
CC were constructed after it was demonstrated that similar improved
CC activities could be achieved using Tie-2 receptor ligands.
CC Angiopoietin-1 (Ang-1) is a Tie-2 receptor ligand and is an agonist for
CC Angiopoietin-2 (Ang-2). Ang-2 is a naturally occurring antagonist
CC of the Tie-2 receptor. The fibrinogen domains (FD) of Ang-1 and Ang-2 are
CC the receptor-binding domains and dimerized versions, of e.g. Ang-1-FD-Fc
CC (Ang-1 fibrinogen domain fused to an Fc domain), can bind to the Tie-2
CC receptor with much higher affinity than monomeric Ang-1-FD (dimerization
CC occurs between the Fc components of adjacent molecules). However,
CC Ang-1-FD-Fc is not able to induce phosphorylation (activate) the Tie-2
CC receptor on endothelial cells unless it is further clustered with goat
CC anti-human Fc antibodies. The novel fusion proteins, mutant versions of
CC Ang-1-FD and Ang-2-FD, were designed that were intrinsically more highly
CC clustered. Tie-2 agonist fusion proteins may be used as haematopoietic
CC factors. Tie-2 receptor antagonist fusion proteins may be used to
CC diagnose or treat, e.g. myeloproliferative or other proliferative
CC disorders of blood forming organs, e.g. thrombocythemia, polycythemia
CC and leukemias.
XX
SQ Sequence 683 AA;

Query Match 89.3%; Score 1277; DB 21; Length 683;
Best Local Similarity 36.5%; Pred. No. 2.2e-35;
Matches 245; Conservative 6; Mismatches 12; Indels 409; Gaps 5;

QY 2 WLKA-----FYDKVAEL----- 14
DB 12 WLVAWVWALCLATPLAKNLEPVSNLNPFLSKGLVIVPKIGDKLDIICPRAEGR 71
QY 15 ----- 14
DB 72 PYEYKLYLVRPEQAACSTVLDPNVLTCTNRPEQEIFRTIKQFSPNMGLEFKKHD 131
QY 15 -----KEAPDLK- 22
DB 132 YYITSTNGSLEGNREGGVCRTTKIIMKVQDPNAVTPQLTTSRPSREADNTVKM 191
QY 23 ----- 22
DB 192 ATQAPGSRGLSDGSKHETVNOEKSGPGASGGSGDPDGFNSKPGKNLEPVSNLSL 251
QY 23 -----AFDKVAEL----- 32
DB 252 NPKFLSGKGLVIVPKIGDKLDIICPRAEAGRPYKYLYLVRPEQAACSTVLDPNVLT 311
QY 33 ----- 32
DB 312 CNRPEQEIFRTIKQFSPNMGLEFKKHDYYITSTNGSLEGNREGGVCRTTKMI 371

QY 33 -----KEA----- 35
DB 372 IMKVQDPNAVTPQLTTSRPSREADNTVKMATQAPGSRGLSDGSKHETVNOEKSGP 431
QY 36 -----FMDKTHTCPPCPAPELILGGPSVFLFPKPKDITLMISR 72
DB 432 GASGGSGDDPDGFNSKPGEPKSCDKTHTCPCPAPELILGGPSVFLFPKPKDITLMISR 491
QY 73 TPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLDHDLN 132
DB 492 TPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLDHDLN 551
QY 133 GKVKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPS 192
DB 552 GKVKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPS 611
QY 193 DIAVEWESNGQPENNYKTTPPVLDSDGSFPLYSKLTVDKSRWQQGNVFSCSVMHEALHNH 252
DB 612 DIAVEWESNGQPENNYKTTPPVLDSDGSFPLYSKLTVDKSRWQQGNVFSCSVMHEALHNH 671
QY 253 YTKSLSLSPGK 264
DB 672 YTKSLSLSPGK 683

RESULT 2
AAB11691
ID AAB11691 standard; Protein; 754 AA.
XX
AC AAB11691;
XX
DT 23-OCT-2000 (first entry)
DE Human secreted Klotho/IgG1 hinge-CH2-CH3 chimeric protein.
XX
KW Human; secreted Klotho/IgG1 heavy chain constant region chimeric protein;
KW immunoglobulin G1; gene therapy; antibody; antiarthritic; nephrotropic;
KW kidney disease; oedema; dropsy; arthritis.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200027885-A1.
XX
PD 18-MAY-2000.
XX
PF 05-NOV-1999; 99WO-JP06152.
XX
PR 05-NOV-1998; 98JP-0314153.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Hanai N, Nakamura K, Kato Y, Nabeshima Y;
XX
DR WPI; 2000-376499/32.
DR N-PSDB; AAA61542, AAA61544.
XX
PT Chimeric polypeptide containing klotho protein bound to an
PT immunoglobulin for treatment and prevention of kidney disease, dropsy
PT and arthritis
XX
PS Claim 9; Page 59-63; 94pp; Japanese.
XX
CC This sequence represents a chimeric protein comprising mature
CC human secreted Klotho protein and the heavy chain constant region
CC (hinge-CH2-CH3) of human IgG1 (immunoglobulin G1). The invention
CC relates to such chimeric proteins comprising the human Klotho protein
CC fused to a human immunoglobulin, and to DNA sequences encoding the chimeric
CC proteins of the invention, to gene therapy vectors comprising DNA encoding the chimeric
CC proteins of the invention, to detection and assay of molecules which
CC interact with Klotho, to antibodies raised against the chimeric protein
CC and to a diagnostic method for Klotho-related disorders using the

CC antibodies. The chimeric proteins of the invention have antiarthritic and
CC nephrotropic activity, and may be used for the treatment, prevention and
CC diagnosis of disorders with which Klotho is associated, including kidney
CC disease, oedema (dropsy) and arthritis.
XX
SQ Sequence 754 AA;

Query Match 89.3%; Score 1277; DB 21; Length 754;
Best Local Similarity 38.0%; Pred. No. 2.8e-35;
Matches 241; Conservative 10; Mismatches 12; Indels 372; Gaps 5;

QY 2 WLKA-----FYDKVAEKLKE----- 16
| : : : : :
Db 120 WARVLPNGSAGVPNREGLYRYRLLERLRELGVQPVVTLVYHWDLPQRLQDAYGQWNRAL 179
QY 17 ----- 16
Db 180 ADHFRDYAELCFRHFQGVQVYITIDNPYVVAWHGYATGRAPGIRGSPRLGYLVVAHNLL 239
QY 17 ----- 16
Db 240 LAHAKVWHLYNTSFRPTGGQGVSTALSSHWINPRMTDHSIKECKSLDFVLGWFAPVF 299
QY 17 ----- 16
Db 300 IDGYPESMKNLSSILPDFTESEKFKIGTADFALCFGLTSLFQLLDPHMKFQLESP 359
QY 17 ----- 16
Db 360 NLQQLSWIDLEFNHPQIFIVENGWVSGTTKRDDAKYMYLKKFIMETLKAIKLDGVDV 419
QY 17 -----FYDKVAEK----- 31
| : : : : :
Db 420 IGYTANSLMDGFEMHRYGYSIRRGIFYVDFLSQDKMLLPKSSALFYQKLIKNGFPPPLPEN 479
QY 32 --LKEAF-----MDKTHTCPPCPAP 49
| : : : : :
Db 480 QLEGTFPCDFAWGVVDNYIQVSQLTKPISSLTKPYHSGSGSGEPKSCDKTHTCPPCPAP 539
QY 50 ELGGPSVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPR 109
| : : : : :
Db 540 ELGGPSVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPR 599
QY 110 EEQNSTYRVSVLTVLHQLDNLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTL 169
| : : : : :
Db 600 EEQNSTYRVSVLTVLHQLDNLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTL 659
QY 170 PSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLT 229
| : : : : :
Db 660 PSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLT 719
QY 230 DKSRWQGNVFCSCVMHEALHNYHTOKSLSPGK 264
| : : : : :
Db 720 DKSRWQGNVFCSCVMHEALHNYHTOKSLSPGK 754

RESULT 3
AAB11693
ID AAB11693 standard; Protein; 787 AA.
XX
AC AAB11693;
XX
XX 23-OCT-2000 (first entry)
XX Human secreted Klotho precursor/IgG1 hinge-CH2-CH3 chimeric protein.
XX Human; secreted Klotho/IgG1 heavy chain constant region chimeric protein;
KW immunoglobulin G1; gene therapy; antibody; antiarthritic; nephrotropic;
KW kidney disease; oedema; dropsy; arthritis.
XX
OS Homo sapiens.
XX Synthetic.
XX

Key Location/Qualifiers
Peptide 1..33
Protein /note= "Signal peptide"
34..787
/note= "Human secreted Klotho/IgG1 heavy chain constant
region chimeric protein (AAB11691)"
WO200027885-A1.
18-MAY-2000.
05-NOV-1999; 99WO-JP06152.
05-NOV-1998; 98JP-0314153.
(KYOW) KYOWA HAKKO KOGYO KK.
Hanai N, Nakamura K, Kato Y, Nabeshima Y;
WPI: 2000-376499/32.
N-PSDB; AAA61544.
Chimeric polypeptide containing klotho protein bound to an
immunoglobulin for treatment and prevention of kidney disease, dropsy
and arthritis
Disclosure; Page 73-79; 94pp; Japanese.
CC This sequence represents a chimeric protein comprising human secreted
CC Klotho protein precursor and the heavy chain constant region
CC (hinge-CH2-CH3) of human IgG1 (immunoglobulin G1). The invention
CC relates to such chimeric proteins comprising the human Klotho protein
CC fused to a human immunoglobulin, and to DNA sequences encoding them. It
CC also relates to gene therapy vectors comprising DNA encoding the chimeric
CC proteins of the invention, to detection and assay of molecules which
CC interact with Klotho, to antibodies raised against the chimeric protein
CC and to a diagnostic method for Klotho-related disorders using the
CC antibodies. The chimeric proteins of the invention have antiarthritic and
CC nephrotropic activity, and may be used for the treatment, prevention and
CC diagnosis of disorders with which Klotho is associated, including kidney
CC disease, oedema (dropsy) and arthritis.
XX
SQ Sequence 787 AA;
Query Match 89.3%; Score 1277; DB 21; Length 787;
Best Local Similarity 38.0%; Pred. No. 3e-35;
Matches 241; Conservative 10; Mismatches 12; Indels 372; Gaps 5;
QY 2 WLKA-----FYDKVAEKLKE----- 16
| : : : : :
Db 153 WARVLPNGSAGVPNREGLYRYRLLERLRELGVQPVVTLVYHWDLPQRLQDAYGQWNRAL 212
QY 17 ----- 16
Db 213 ADHFRDYAELCFRHFQGVQVYITIDNPYVVAWHGYATGRAPGIRGSPRLGYLVVAHNLL 272
QY 17 ----- 16
Db 273 LAHAKVWHLYNTSFRPTGGQGVSTALSSHWINPRMTDHSIKECKSLDFVLGWFAPVF 332
QY 17 ----- 16
Db 333 IDGYPESMKNLSSILPDFTESEKFKIGTADFALCFGLTSLFQLLDPHMKFQLESP 392
QY 17 ----- 16
Db 393 NLQQLSWIDLEFNHPQIFIVENGWVSGTTKRDDAKYMYLKKFIMETLKAIKLDGVDV 452
QY 17 -----AFDWLKA-----FYDKVAEK----- 31
| : : : : :
Db 453 IGYTANSLMDGFEMHRYGYSIRRGIFYVDFLSQDKMLLPKSSALFYQKLIKNGFPPPLPEN 512
QY 32 --LKEAF-----MDKTHTCPPCPAP 49

Db 513 QPLEGTFPCDFAWGVNDVNIQVSQLTKPISSLTQPKYHGGSGGSEKSCDKTHTCPCCPAP 572
QY 50 ELGGSPVFLFPKPKDMLISRTPEVTCVVVDVSHEDPEVKENWYVDGVEVHNATKPR 109
Db 573 ELGGSPVFLFPKPKDMLISRTPEVTCVVVDVSHEDPEVKENWYVDGVEVHNATKPR 632
QY 110 EEQNSTYRVVSVLTLDHDLNGKEYCKKVSNNKALPAPIEKTISKAKGQPREPQVYTLTP 169
Db 633 EEQNSTYRVVSVLTLDHDLNGKEYCKKVSNNKALPAPIEKTISKAKGQPREPQVYTLTP 692
QY 170 PSRDELTKNQVSTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFPLYSKLTV 229
Db 693 PSRDELTKNQVSTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFPLYSKLTV 752
QY 230 DKSRWQOGNVFSCVMHEALHNYTKQSLSPGK 264
Db 753 DKSRWQOGNVFSCVMHEALHNYTKQSLSPGK 787
RESULT 4
ID AAW70539 standard; Protein; 1218 AA.
XX AAW70539;
DT 26-JAN-1999 (first entry)
DE Integrin alpha-4 chain.
KW Integrin; alpha-4 chain; immunoglobulin; chimeric; heterodimer complex;
KW inhibitor; binding; ligand; blood platelet; hemostatic; diagnostic agent;
KW human.
OS Homo sapiens.
FH Key
FT Peptide 1..39 Location/Qualifiers
FT Protein 40..1218 /note= "signal peptide"
FT Protein /note= "mature protein"
XX WO9832771-Al.
XX 30-JUL-1998.
XX 29-JAN-1998; 98WO-JP00370.
XX 29-AUG-1997; 97JP-0234544.
XX 29-JAN-1997; 97JP-0015118.
XX (TORA) TORAY IND INC.
XX Kainoh M, Tanaka T;
XX WPI; 1998-427881/36.
XX N-PSDB; AAV33772.
XX Integrin-immunoglobulin chimeric protein heterodimer complexes as
PT platelet substitutes - contain the alpha and beta integrin chains
PT associated in stable state and bind to extracellular matrix in the
PT presence of plasma components
XX Claim 7; Pages 42-50; 87pp; Japanese.
XX This represents an integrin alpha-4 chain. The invention provides
CC integrin-immunoglobulin chimeric protein heterodimer complexes that
CC comprise an integrin alpha or beta chain associated with an
CC immunoglobulin light or heavy chain. These chimeric proteins form
CC heterodimer complexes, in particular with a chimeric protein containing
CC an integrin alpha chain and an immunoglobulin chain with a chimeric
CC protein containing an integrin beta chain and an immunoglobulin chain;
CC the immunoglobulin chain in each case may be a heavy chain, or one of the

CC two may be a light chain. The integrin alpha chain is preferably alpha 4
or alpha 2 and the integrin beta chain is preferably beta 1. Animal cells
CC transformed with vectors containing the DNA coding for the above chimeric
proteins can be used in the preparation of the chimeric proteins and
CC their heterodimer complexes. The heterodimer complexes, which are useful
for testing potential promoters and inhibitors of the binding of
CC integrins to their ligands, function as blood platelet substitutes and
hemostatics and as diagnostic agents.
XX
SQ Sequence 1218 AA;
Query Match 88.9%; Score 1271; DB 19; Length 1218;
Best Local Similarity 24.9%; Pred. No. 1.3e-34;
Matches 243; Conservative 4; Mismatches 14; Indels 716; Gaps 4;
QY 4 KAFYDK----- 9
Db 242 KAFLDKQNVKFGSYLGYSVGAGHFRSQHTTEVVGAPQHEQIGKAYIFSIDEKELNHL 301
QY 10 ----- 9
Db 302 EMKGKLGYSFGASCAVDLNADGFSDDLVLGAPMQSTIREGRVYVYINSGGAVMNAME 361
QY 10 ----- 9
Db 362 TNLVGSKYAARFGESIVNLGIDIDNGDFEDVAIGAPQEDDLQGAIIYINGRADGISSTFS 421
QY 10 ----- 9
Db 422 QRIEGLQISKSLSMFGQSIGQIDADNNGYVDVAVGAFRSDSAVLLTRPPVIVDASLSH 481
QY 10 ----- 9
Db 482 PESVNRKFCDCVENGWPSVICDLTLCSYKGVKPGYIVLIFYNMSLDVNRKABSPREFY 541
QY 10 ----- 9
Db 542 SSGTSDVITGSTQVSSREANCRTHOAFMRKDVDRDILTPTQIEAAVHLGPHVSKRSTEE 601
QY 10 ----- 9
Db 602 FPLPQILOQKKEKDIMKKTINFARFCAHENSADLQVSAKIGFLAPHENKTYLAVGSMK 661
QY 10 ----- 9
Db 662 TMLNVSLFNAGDDAYETTLHVKLPGVLYFIKILEEEKQINCEVTDNSGVQLDCSIGY 721
QY 10 ----- 9
Db 722 IYVDHLRIDISFLLDVSSLSRAEEDLSITVHATCENEEMDNKHSRVTVAIPLKYEYK 781
QY 10 ----- 9
Db 782 LTVHGFVNPTSFYGSNDENEPETCMVKMNLTFHVINTGNSMAPNVSVIWPVNSFSPQ 841
QY 10 -----VAEKLKEAFOWLKA---FYDKVAEKL----- 32
Db 842 TDKLFNLVDVQTGTGECHEFNQVRCALQEQKSAQMTLQAGIVRFLSKTDKRLLYCIKADP 901
QY 33 -----KEAFM----- 37
Db 902 HCLFLNCLNFGKMSGKASVHIQLEGRPSILEMDETSALKFEIRATGFPENPRVIELNK 961
QY 38 -----DKTHCTPCPAPALGGLGGSVFLFPPKPKDT 67
Db 962 DENVAHVLEGLHQRKRYFTDPEEPKSCDKTHTCPAPALGGLGGSVFLFPPKPKDT 1021
QY 68 LMISRTPEVTCVVVDVSHEDPEVKENWYVDGVEVHNATKPREEOYNSYRVVSVLTVLH 127
Db 1022 LMISRTPEVTCVVVDVSHEDPEVKENWYVDGVEVHNATKPREEOYNSYRVVSVLTVLH 1081
QY 128 QDWLNGKEYCKKVSNNKALPAPIEKTISKAKGQPREPQVYTLTPSRDELTKNQVSLTCLVK 187
|||||


```

QY 38 ----- 37
Db 1081 NGTFASSTQVLTAAAEINTYNYPIYVIEDNTVTIPLMINKPDEKAEVPTDPEPKSC 1140
QY 38 DKTHTCPCPCAPPELLGGPSVLEFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 97
Db 1141 DKTHTCPCPCAPPELLGGPSVLEFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 1200
QY 98 GVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTSKAK 157
Db 1201 GVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTSKAK 1260
QY 158 GQPREQVYTLPPSRDELTKNQVSLTCLVKGYFSPVDIAVWESNGQPENNYKTTPPVLD 217
Db 1261 GQPREQVYTLPPSRDELTKNQVSLTCLVKGYFSPVDIAVWESNGQPENNYKTTPPVLD 1320
QY 218 DGSFFLYSKLTVDKSRWQOGNPFSCVMHEALHNHYTQKSLSLSPGK 264
Db 1321 DGSFFLYSKLTVDKSRWQOGNPFSCVMHEALHNHYTQKSLSLSPGK 1367

RESULT 6
AAU09817
ID AAU09817 standard; Protein; 594 AA.
XX
AC AAU09817;
XX
DT 27-FEB-2002 (first entry)
XX
DE Murine FGFR-L extracellular domain-Fc fusion protein.
XX
KW Mouse; fibroblast growth factor receptor-like protein; FGFR-L; anorectic;
KW haemostatic; osteopathic; cytostatic; nephrotropic; antidiabetic;
KW immunomodulator; antiinflammatory; haematopoietic disorder; osteoporosis;
KW osteogenesis imperfecta; Paget's disease; periodontal disease; cancer;
KW hypercalcaemia; acute glomerulonephritis; chronic glomerulonephritis;
KW diabetes; obesity; cachexia; transgenic animal; gene therapy; ds.
XX
OS Mus musculus.
XX
FH Key
FT 1..1785
FT CDS
FT /tag= a
FT /product= "Fibroblast growth factor receptor-like protein
FT extracellular domain-Fc fusion protein"
XX
PN WO200170977-A2.
XX
PD 27-SEP-2001.
XX
PF 22-MAR-2001; 2001WO-US09073.
XX
PR 22-MAR-2000; 2000US-191379P.
XX
PA (AMGE-) AMGEN INC.
PA (SARI/) SARIS C M.
PA (MUSX/) MU S X.
PA (XIAM/) XIAM M.
PA (BOON/) BOONE T C.
PA (COVE/) COVEY T.
XX
PI Saris CM, Mu SX, Xia M, Boone TC, Covey T;
XX
DR WPI; 2001-626128/72.
XX
DR N-PSDB; AAS14939.
XX
PT Novel nucleic acid encoding fibroblast growth factor receptor-like
PT polypeptides, useful for treating hematopoietic disorder, osteoporosis,
PT Paget's disease, glomerulonephritis, cancer, diabetes, obesity and
PT cachexia
XX
PS Example 4; Page 161-163; 163pp; English.

```

```

XX The invention relates to a novel isolated fibroblast growth factor
CC receptor-like (FGFR-L) polypeptide (I). (I) and the nucleic acid (II)
CC encoding (I) are useful for treating, preventing or ameliorating
CC a medical condition including haematopoietic disorder, osteoporosis,
CC osteogenesis imperfecta, Paget's disease, periodontal disease,
CC hypercalcaemia, acute glomerulonephritis, chronic glomerulonephritis,
CC cancer, diabetes, obesity and cachexia. (I) is also useful for
CC identifying a compound which binds to FGFR-L polypeptide, by contacting
CC (I) with a compound, determining the extent of binding of the FGFR-L
CC polypeptide to the compound, and determining the activity of the
CC polypeptide when bound to the compound. (II) is useful for modulating
CC levels of a polypeptide in an animal. A transgenic animal comprising (I)
CC is useful for determining whether a compound inhibits FGFR-L polypeptide
CC activity or FGFR-L polypeptide production, by exposing the transgenic
CC animal to the compound and measuring FGFR-L polypeptide or production in
CC the animal. (II) is useful for mapping the locations of FGFR-L gene and
CC related genes on chromosomes, as hybridisation probes in diagnostic
CC assays to test for the presence of an FGFR-L nucleic acid molecule in
CC mammalian tissue or bodily fluid samples, in gene therapy, and as tools
CC for isolating corresponding FGFR-L polypeptide genes. (I) is useful as
CC immunogen, and for cloning FGFR-L polypeptide ligands using an
CC expression cloning strategy. The present sequence represents the
CC coding sequence of mouse fibroblast growth factor receptor-like protein
CC extracellular domain-Fc fusion protein used to generate anti-FGFR-L
CC antibodies as described in the invention.
XX
SQ Sequence 594 AA;
Query Match 88.7%; Score 1268; DB 22; Length 594;
Best Local Similarity 57.6%; Pred. No. 3.3e-35;
Matches 239; Conservative 9; Mismatches 15; Indels 152; Gaps 4;
QY 2 WLK----- 4
Db 180 WMKDDQTLTHLEASERKKKWTLSLKNLKPDSGGYTCRVSNKAGAINATYKVDVIQRT 239
QY 5 -----AFYDKVAEKLKEAFDWLK----- 22
Db 240 SKPVLGTGTHPVNTTVDGTTSTFOCKVRSDVKPVIQWLKRVYSGEGRHNSTIDVGQKF 299
QY 23 -----AFYDK-----VAEKLKEAFM----- 37
Db 300 VVLPTCDVNSRDGYSYLNKLLISRARQDDAGMYICLGANTMGYSRSPAFLTLPDPKPPG 359
QY 38 -----DKHTCCPCAPPELLGGPSVLEFPKPKDTLMISRTPEVTCVVVDVSHEDPE 89
Db 360 PPMASSVDKTHTCPCAPPELLGGPSVLEFPKPKDTLMISRTPEVTCVVVDVSHEDPE 419
QY 90 VKFNWYVDGVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPI 149
Db 420 VKFNWYVDGVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPI 479
QY 150 EKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGYFSPVDIAVWESNGQPENNYK 209
Db 480 EKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGYFSPVDIAVWESNGQPENNYK 539
QY 210 TTPPVLDSDGSFFLYSKLTVDKSRWQOGNPFSCVMHEALHNHYTQKSLSLSPGK 264
Db 540 TTPPVLDSDGSFFLYSKLTVDKSRWQOGNPFSCVMHEALHNHYTQKSLSLSPGK 594

RESULT 7
AAW07098
ID AAW07098 standard; protein; 951 AA.
XX
AC AAW07098;
XX
DT 03-FEB-1999 (first entry)
XX
DE Human gp130-C-gamma-1 amino acid sequence.
XX
KW gp130; cytokine antagonist; Interleukin; gamma-interferon.

```


KW granulocyte macrophage colony-stimulating factor; J peptide;
XX transforming growth factor-beta.
OS Synthetic.
XX Homo sapiens.
FH Key Location/Qualifiers
FT Protein 1..619
FT /note= "human gp130"
FT Protein 662..651
FT /note= "from the constant region of human IgG1"
XX US5844099-A.
XX 01-DEC-1998.
XX 27-NOV-1995; 95US-0563105.
XX 27-NOV-1995; 95US-0563105.
PR 20-OCT-1993; 93US-0140222.
XX (REGE-) REGENERON PHARM INC.
XX Economides A, Stahl N, Yancopoulos GD;
XX WPI; 1999-044669/04.
XX Cytokine antagonists - comprising extracellular domains of
PT specificity-determining and signal-transducing components of
PT cytokine receptor
XX Example 4; Fig 9A-B; 46pp; English.
XX The present sequence represents the amino acid sequence of human
CC gp130-C-gamma-1. The protein is used in the course of the invention.
CC The specification describes cytokine antagonists comprising only the
CC extracellular domain of the specificity-determining component of
CC the cytokine receptor and the extracellular domain of a
CC signal-transducing component of the cytokine receptor. The cytokine
CC is an interleukin (IL-1, IL-2, IL-3, IL-4, IL-5 or IL-15),
CC granulocyte macrophage colony-stimulating factor (GM-CSF),
CC gamma-interferon or transforming growth factor-beta (TGF-beta). The
CC antagonist is capable of binding the cytokine to form a nonfunctional
CC complex. The compounds have therapeutic activity as cytokine antagonists
CC and can also be used in assays for identifying novel agonists and
CC antagonists of cytokines.
XX Sequence 951 AA;
SQ
Query Match 88.7%; Score 1268; DB 20; Length 951;
Best Local Similarity 38.4%; Pred. No. 9.5e-35;
Matches 246; Conservative 8; Mismatches 10; Indels 376; Gaps 7;
QY 1 DW----- 2
DB 312 DWSEASGITYEDRPSKAPFWYKIDPSHTQGYRTVQVWVMTLPPFEANGKILDYEVTLT 371
QY 3 -----LK 4
DB 372 RWKSHLQYTVNATKLTNLTNDRYLATLVNRLVGKSDAAVLTPACDFQATHPVMDLK 431
QY 5 AF-----YDK-----VAEK-- 13
DB 432 AFPKDNMLVWETTPRESVKYILEWCVLSDKAPCIDTQWQEDGTVHRTVLRGNLAESKC 491
QY 14 -----LKEA-----FDW----- 20
DB 492 YLITVTPYADPGSPESIRAYLKQAPPSKGTVRTTKVGKNEAVLEMDQLPVDVQNGFI 551
QY 21 -----LKAFYD----- 26
DB 552 RNVTIFYRTIIGNETAVNVDSHSHTPYTLSSLTSDTLVWVRMAAYTDEGKGDPFTTTP 611

QY 27 ----- 26
DB 612 KFAGEIESGASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPTVTSWNSGALISGV 671
QY 27 -----KVAELKLEAFMDKTHTCP 44
DB 672 HTFPAVLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCP 731
QY 45 PCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNA 104
DB 732 PCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNA 791
QY 105 KTKPREEQYNSTYRWVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKAKGQPREPQ 164
DB 792 KTKPREEQYNSTYRWVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKAKGQPREPQ 851
QY 165 VYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFELY 224
DB 852 VYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFELY 911
QY 225 SKLTVDKSRWQOGNVSFCSVMHEALHNHYTOKLSLSPGK 264
DB 912 SKLTVDKSRWQOGNVSFCSVMHEALHNHYTOKLSLSPGK 951
RESULT 8
AA92186
ID AA92186 standard; protein; 951 AA.
XX AC AA92186;
XX DT 01-AUG-2000 (first entry)
XX DE Human gp130-C-gamma-1.
XX KW gp130-C-gamma-1; cytokine; antagonist; CNTF; receptor; fusion protein;
XX cytotostatic; immunomodulator; osteopathic.
XX OS Synthetic.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Protein 1..619
FT /label= gp130
FT Peptide 620..621
FT /note= "Ser-Gly bridge"
FT Protein 622..951
FT /label= IgG1_constant_region
XX WO200018932-A2.
XX PN 06-APR-2000.
XX PD 22-SEP-1999; 99WO-US22045.
XX PF 25-SEP-1998; 98US-0101858.
XX PR 19-MAY-1999; 99US-0313942.
XX PA (REGE-) REGENERON PHARM INC.
XX PI Stahl N, Yancopoulos GD;
XX PI WPI; 2000-293165/25.
XX DR Isolated nucleic acid molecule for treating cytokine-related diseases
XX or disorders encodes a fusion polypeptide capable of binding a cytokine
XX to form a nonfunctional complex
XX Example 4; Fig 9; 152pp; English.
XX The invention concerns production of antagonists to any cytokine that
XX utilizes an alpha specificity determining component, which when combined
XX with the cytokine, binds to a first beta signal transducing component to

CC form a non-functional intermediate which then binds to a second beta
CC signal transducing component causing beta-receptor dimerization, the
CC soluble alpha specificity determining component of the receptor,
CC (SR-alpha) and the extracellular domain of the first beta signal
CC transducing component of the cytokine receptor (beta-1) are combined to
CC form heterodimers (SR-alpha:beta-1) that act as antagonist to the
CC cytokine by binding the cytokine to form a non-functional complex. The
CC receptor components are shared by cytokines such as the CNTF (ciliary
CC neurotrophic factor) family of cytokines. The invention provides the
CC basis for the development of IL-6 antagonists, as they show that if, in
CC the presence of a ligand, a non-functional intermediate complex,
CC consisting of the ligand, its alpha receptor and its beta-1 receptor
CC component, can be formed, it will effectively block the action of the
CC ligand. Effective antagonists of IL-6 or CNTF consist of heterodimers
CC of the extracellular domains of the alpha specificity determining
CC components of their receptors and the extracellular domain of gp130.
CC The resultant heterodimers, function as high-affinity traps, rendering
CC the cytokine inaccessible to form a signal transducing complex with the
CC native membrane-bound forms of their receptor. The nucleic acids and
CC polypeptides are useful for treating cytokine-related diseases or
CC disorders such as osteoporosis and primary and secondary effects of
CC cancer including multiple myeloma or cachexia.

xx SQ Sequence 951 AA;

Query Match 88.7%; Score 1268; DB 21; Length 951;
Best Local Similarity 38.4%; Pred No. 9.5e-35;
Matches 246; Conservative 8; Mismatches 10; Indels 376; Gaps 7;

QY 1 DW-----YDK--VAEK-- 13
||
Db 312 DWSEASGITYEDRPSKAPFWKIDPSHTQGYRTVQLVWKVTLPPPEANGKILDYEVTLT 371
QY 3 -----LK 4
||
Db 372 RKMSHLQNTVNAKLTVALNTLNDRLATLTVRLNKGSDAAVLITPACDFQATHPVMDLK 431
QY 5 AF-----YDK--VAEK-- 13
||
Db 432 AFPKNMLWVWETTPRESVKYKYLEWCVLSDKAPCIDWQOEDGIVHRYTLRGNLAESKC 491
QY 14 -----LKEA-----FDW----- 20
||
Db 492 YLITVTPVYADGPGSEIKAYLKQAPPSKGTVRTKVKGNKAEVLEWDLQVPDVQNGFI 551
QY 21 -----LKAFYD----- 26
||
Db 552 RNYTFYRTIIGNETAVNVDSSTHTYTLSSLTSDTLYMVVMAAYTDEGCKGPEFTFTTP 611
QY 27 ----- 26
Db 612 KFAQGEIESGASTKGPSVFPPLAPSKSTSGGTAALGCLVKDYFPPEVTVSWNSGALTSV 671
QY 27 -----KVAEKLKEAFMDKTHTCP 44
||
Db 672 HTFPVAVLQSSGLYSLSSVVTPSSISLGTQYICNVNHNKPSNFKDKVPEKSCDKTHTCP 731
QY 45 PCPAPPELLGGSVFLFPKPKDTLMSLRPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNA 104
||
Db 732 PCPAPPELLGGSVFLFPKPKDTLMSLRPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNA 791
QY 105 KTKPREEQYNTRYVSVLTVLHQDLNKGKEYKCKVSKNKAALPAPTEKTSKAKGQPREPQ 164
||
Db 792 KTKPREEQYNTRYVSVLTVLHQDLNKGKEYKCKVSKNKAALPAPTEKTSKAKGQPREPQ 851
QY 165 VYTLPPSRDELTKNOVSLTCLVKGYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLY 224
||
Db 852 VYTLPPSRDELTKNOVSLTCLVKGYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLY 911
QY 225 SKLTVDKSRWQOGNFGVSCSVMEALHNHYTQKSLSLSPCK 264
||
Db 912 SKLTVDKSRWQOGNFGVSCSVMEALHNHYTQKSLSLSPCK 951

RESULT 9
AAY92187
ID AAY92187 standard; protein; 961 AA.
XX
AC AAY92187;
XX
DT 01-AUG-2000 (first entry)
XX
DE Human gp130-J-C-gamma-1.
XX
KW gp130-J-C-gamma-1; cytokine; antagonist; CNTF; receptor; fusion protein;
KW cytostatic; immunomodulator; osteopathic.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Protein 1..619
FT /label= gp130
FT Peptide 620..621
FT /note= "Ser-Gly bridge"
FT Peptide 622..631
FT /note= "J-peptide"
FT Protein 632..961
FT /label= IgG1_constant_region
XX
PN WO200018932-A2.
XX
PD 06-APR-2000.
XX
PF 22-SEP-1999; 99WO-US22045.
XX
PR 25-SEP-1998; 98US-0101858.
PR 19-MAY-1999; 99US-0313942.
XX
PA (REGE-) REGENERON PHARM INC.
XX
PI Stahl N, Yancopoulos GD;
XX
DR WPI; 2000-293165/25.
XX
PT Isolated nucleic acid molecule for treating cytokine-related diseases
PT or disorders encodes a fusion polypeptide capable of binding a cytokine
PT to form a nonfunctional complex
XX
PS Example 4; Page -: 152pp; English.
XX
CC The invention concerns production of antagonists to any cytokine that
CC utilizes an alpha specificity determining component, which when combined
CC with the cytokine, binds to a first beta signal transducing component to
CC form a non-functional intermediate which then binds to a second beta
CC signal transducing component causing beta-receptor dimerization, the
CC soluble alpha specificity determining component of the receptor
CC (SR-alpha) and the extracellular domain of the first beta signal
CC transducing component of the cytokine receptor (beta-1) are combined to
CC form heterodimers (SR-alpha:beta-1) that act as antagonist to the
CC cytokine by binding the cytokine to form a non-functional complex. The
CC receptor components are shared by cytokines such as the CNTF (ciliary
CC neurotrophic factor) family of cytokines. The invention provides the
CC basis for the development of IL-6 antagonists, as they show that if, in
CC the presence of a ligand, a non-functional intermediate complex,
CC consisting of the ligand, its alpha receptor and its beta-1 receptor
CC component, can be formed, it will effectively block the action of the
CC ligand. Effective antagonists of IL-6 or CNTF consist of heterodimers
CC of the extracellular domains of the alpha specificity determining
CC components of their receptors and the extracellular domain of gp130.
CC The resultant heterodimers, function as high-affinity traps, rendering
CC the cytokine inaccessible to form a signal transducing complex with the
CC native membrane-bound forms of their receptor. The nucleic acids and
CC polypeptides are useful for treating cytokine-related diseases or
CC disorders such as osteoporosis and primary and secondary effects of
CC cancer including multiple myeloma or cachexia.

```

XX SQ Sequence 961 AA;
Query Match 88.7%; Score 1268; DB 21; Length 961;
Best Local Similarity 37.1%; Pred. No. 9.7e-35;
Matches 246; Conservative 8; Mismatches 10; Indels 386; Gaps 7;

Qy 1 DW----- 2
Db 312 DMSEASGITYEDRPSKAPFWYKIDPSHTQGYRTVQLVWKTLPFPFANGKILDEYVTLT 371
Qy 3 -----LK 4
Db 372 RWKSHLQNTVYATKLTATVNLNDRLATLTVRNLVKGSDAAVLTIPACDFQATHPVMDLK 431
Qy 5 AF-----YDK-----VAEK-- 13
Db 432 AFPPDNMLVWETTPRESVKYILEWCVLSDKAPCIDTWOQEDGTVHRTYLRGNLAESKC 491
Qy 14 -----LKEA-----FDW----- 20
Db 492 YLITVTPVYADPGSPESIKAYLQAPSGKPTVTRTKVGNKAEVLEWDQLPVDVQNGFI 551
Qy 21 -----LKAFYD----- 26
Db 552 RNYTIFYRTIIGNETAVNVDSSHTYLLSLTSDTLVYMRMAAYTDEGGKDGPEFTFTTP 611
Qy 27 ----- 26
Db 612 KFAQGEIESGGQTLVTVSSASTKPSVFPPLAPSSKSTSGTAALGCLVLYDFPEPVTVS 671
Qy 27 -----KVAELKE 34
Db 672 WNSGALTSQVHTFPAVLQSSGLYSLSSVTVVPSSSLGTOTVICNVNHNKPSNTKVDKKEP 731
Qy 35 AFMDKTHFCPCPAPELGGPSVLFPPKPKDTLMISRTPEVTVVVDVSHEDPEVKFNW 94
Db 732 KSCDKTHFCPCPAPELGGPSVLFPPKPKDTLMISRTPEVTVVVDVSHEDPEVKFNW 791
Qy 95 YVDGVEVHNATKPREQYNTYRVSVVLTVLHODWLNKGYCKVSNKALPAPTEKTIS 154
Db 792 YVDGVEVHNATKPREQYNTYRVSVVLTVLHODWLNKGYCKVSNKALPAPTEKTIS 851
Qy 155 KAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPYPSDIAVEWESNGQPPENNYKTPPV 214
Db 852 KAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFPYPSDIAVEWESNGQPPENNYKTPPV 911
Qy 215 LDSGSPFLYKLTVDKSRWQGNVFCVSNMHEALHNNHYTKSLSPGK 264
Db 912 LDSGSPFLYKLTVDKSRWQGNVFCVSNMHEALHNNHYTKSLSPGK 961
```

```

RESULT 10
AA92205
ID AA92205 standard; Protein; 1158 AA.
XX AC
XX AC AA92205;
XX DT 01-AUG-2000 (first entry)
XX DE
XX DE Fusion polypeptide 616, IL-6 trap.
XX KW IL-6 trap; cytokine; antagonist; CNTF; receptor; fusion protein;
XX KW cytosolic; immunomodulator; osteopathic.
XX OS Synthetic.
XX OS Homo sapiens.
XX PN WO200018932-A2.
XX PD 06-APR-2000.
XX PF 22-SEP-1999; 99WO-US22045.
```

```

XX PR 25-SEP-1998; 98US-0101858.
XX PR 19-MAY-1999; 99US-0313942.
XX PA (REGE-) REGENERON PHARM INC.
XX PI Stahl N, Yancopoulos GD;
XX PI WPI: 2000-293165/25.
XX DR N-PSDB; AAA09047.
XX DR
XX PT Isolated nucleic acid molecule for treating cytokine-related diseases
XX PT or disorders encodes a fusion polypeptide capable of binding a cytokine
XX PT to form a nonfunctional complex
XX PS Example 6; Fig 25A-F; 152pp; English.
XX CC This sequence shows fusion polypeptide 616, which is capable of
XX CC binding cytokine IL-6 to form a non-functional complex.
XX CC The invention concerns production of antagonists to any cytokine that
XX CC utilizes an alpha specificity determining component, which when combined
XX CC with the cytokine, binds to a first beta signal transducing component to
XX CC form a non-functional intermediate which then binds to a second beta
XX CC signal transducing component causing beta-receptor dimerization, the
XX CC soluble alpha specificity determining component of the receptor
XX CC (SR-alpha) and the extracellular domain of the first beta signal
XX CC transducing component of the cytokine receptor (beta-1) are combined to
XX CC form heterodimers (SR-alpha:beta-1) that act as antagonist to the
XX CC cytokine by binding the cytokine to form a non-functional complex. The
XX CC receptor components are shared by cytokines such as the CNTF (cllary
XX CC neurotrophic factor) family of cytokines. The invention provides the
XX CC basis for the development of IL-6 antagonists, as they show that if, in
XX CC the presence of a ligand, a non-functional intermediate complex,
XX CC consisting of the ligand, its alpha receptor and its beta-1 receptor
XX CC component, can be formed, it will effectively block the action of the
XX CC ligand. Effective antagonists of IL-6 or CNTF consist of heterodimers
XX CC of the extracellular domains of the alpha specificity determining
XX CC components of their receptors and the extracellular domain of gp130.
XX CC The resultant heterodimers, function as high-affinity traps, rendering
XX CC the cytokine inaccessible to form a signal transducing complex with the
XX CC native membrane-bound forms of their receptor. The nucleic acids and
XX CC polypeptides are useful for treating cytokine-related diseases or
XX CC disorders such as osteoporosis and primary and secondary effects of
XX CC cancer including multiple myeloma or cachexia.
XX SQ Sequence 1158 AA;
```

```

Query Match 88.7%; Score 1268; DB 21; Length 1158;
Best Local Similarity 37.1%; Pred. No. 1.5e-34;
Matches 242; Conservative 12; Mismatches 9; Indels 390; Gaps 6;

Qy 2 WLKA-----FYDKV-----AEKL----- 14
Db 506 WVEAENALGKVTSDHINFDPVYKVPNPPHNLVSNSELSEILKLTWTNPSIKSVILK 565
Qy 15 -----KEAFDW----- 20
Db 566 YNIQRTKDASTWSQIIPEDTASTRSSFTVQDLKPFTEYVPIRCMKEDGKGYNDWSEE 625
Qy 21 ----- 20
Db 636 ASGITVEDRPSKAPSWYKIDPSHTQGYRTVQLVWKTLPFPFANGKILDEYVTLTRKSH 685
Qy 21 -----LKAF--- 24
Db 686 LQNTVYATKLTATVNLNDRLATLTVRNLVKGSDAAVLTIPACDFQATHPVMDLKAFPKD 745
Qy 25 ----- 24
Db 746 NMLVWETTPRESVKYILEWCVLSDKAPCIDTWOQEDGTVHRTYLRGNLAESKYLITV 805
Qy 25 -----YDKVAELKEAFM----- 37
```

Db 806 TPVYADGPGSPESIKAYLKOAPPSKGTPTVTKKVGKNEAVLEWDQLPVDVQNGFIRNYTI 865
QY 38 -----DKTHTCTPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 37
Db 866 FYRTIIGNEAVNVDSSTHTYTLSSLTSDLYMVRMAAYTDEGKGDPETFTTPKFAQG 925
QY 38 -----DKTHTCTPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 91
Db 926 EIESGGDKTHTCTPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 985
QY 92 FNMVYDGVVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKKVSNAKALPAPIEK 151
Db 986 FNMVYDGVVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKKVSNAKALPAPIEK 1045
QY 152 TISKAKGQPREPOVYITLPPSRDELTKNQVSLTCLVKGFYPSDIAVEHESNGQPENNYKTT 211
Db 1046 TISKAKGQPREPOVYITLPPSRDELTKNQVSLTCLVKGFYPSDIAVEHESNGQPENNYKTT 1105
QY 212 PPVLDSGGSFFLYSKLTVDSKRWQGNVFCVSMHEALHNHYTKLSLSPGK 264
Db 1106 PPVLDSGGSFFLYSKLTVDSKRWQGNVFCVSMHEALHNHYTKLSLSPGK 1158

RESULT 11
AAY92204
ID AAY92204 standard; Protein; 1168 AA.
AC AAY92204;
CX
DT 01-AUG-2000 (first entry)
DE Fusion polypeptide 412, IL-6 trap.
DE
KW IL-6 trap; cytokine; antagonist; CNTF; receptor; fusion protein;
KW cytostatic; immunomodulator; osteopathic.
XX
OS Synthetic.
OS Homo sapiens.
PN WO200018932-A2.
PD 06-APR-2000.
PF 22-SEP-1999; 99WO-US22045.
PR 25-SEP-1998; 98US-0101858.
PR 19-MAY-1999; 99US-0313942.
XX
PA (REG-) REGENERON PHARM INC.
PI Stahl N, Yancopoulos GD;
XX
XX WPI: 2000-293165/25.
DR N-PSDB; AAA09046.
XX
PT Isolated nucleic acid molecule for treating cytokine-related diseases
PT or disorders encodes a fusion polypeptide capable of binding a cytokine
PT to form a nonfunctional complex
XX
PS Example 6; Fig 24A-F; 152pp; English.
XX
CC This sequence shows fusion polypeptide 412, which is capable of
CC binding cytokine IL-6 to form a non-functional complex.
CC The invention concerns production of antagonists to any cytokine that
CC utilizes an alpha specificity determining component, which when combined
CC with the cytokine, binds to a first beta signal transducing component to
CC form a non-functional intermediate which then binds to a second beta
CC signal transducing component causing beta-receptor dimerization, the
CC soluble alpha specificity determining component of the receptor
CC (sr-alpha) and the extracellular domain of the first beta signal
CC transducing component of the cytokine receptor (beta-1) are combined to
CC form heterodimers (sr-alpha:beta-1) that act as antagonist to the
CC cytokine by binding the cytokine to form a non-functional complex. The

CC receptor components are shared by cytokines such as the CNTF (ciliary
CC neurotrophic factor) family of cytokines. The invention provides the
CC basis for the development of IL-6 antagonists, as they show that if, in
CC the presence of a ligand, a non-functional intermediate complex,
CC consisting of the ligand, its alpha receptor and its beta-1 receptor
CC component, can be formed, it will effectively block the action of the
CC ligand. Effective antagonists of IL-6 or CNTF consist of heterodimers
CC of the extracellular domains of the alpha specificity determining
CC components of their receptors and the extracellular domain of gp130.
CC The resultant heterodimers, function as high-affinity traps, rendering
CC the cytokine inaccessible to form a signal transducing complex with the
CC native membrane-bound forms of their receptor. The nucleic acids and
CC polypeptides are useful for treating cytokine-related diseases or
CC disorders such as osteoporosis and primary and secondary effects of
CC cancer including multiple myeloma or cachexia.
XX
SQ Sequence 1168 AA;
Query Match 88.7%; Score 1268; DB 21; Length 1168;
Best Local Similarity 37.1%; Pred. No. 1.5e-34;
Matches 242; Conservative 12; Mismatches 9; Indels 390; Gaps 6;
QY 2 WLKA-----FYDKV-----AEKL----- 14
Db 516 WVEAENALGKVTSDHINFDPPYKVPKPNPHNLSVINSEELSLKLTWTNPSIKSVILK 575
QY 15 -----KEAFDW----- 20
Db 576 YNIOYRTKDASTWSQIPPEDTASTRSFTVODLKPTEYVFRICMKEDGKGYWSDSEE 635
QY 21 ----- 20
Db 636 ASGITVEDRPSKAPSWYKIDPSHTOGYRTVQLVWKTLPPEFANGKILDYEVTLRWKSH 695
QY 21 -----LKAF--- 24
Db 696 LQNYTVNATKLTNLTNDRYLATLTVRNLVGKSDAAVLTIPACDFQATHPVMCLKAFPKD 755
QY 25 ----- 24
Db 756 NMLWVETTPRESVKKYILEWCVLSDKAPCITDMQEDGTVHRTYLRGNLAESKCYLITV 815
QY 25 -----YDKVAEKLKEAFM----- 37
Db 816 TPVYADGPGSPESIKAYLKOAPPSKGTPTVTKKVGKNEAVLEWDQLPVDVQNGFIRNYTI 875
QY 38 ----- 37
Db 876 FYRTIIGNETAVNVDSSTHTYTLSSLTSDLYMVRMAAYTDEGKGDPETFTTPKFAQG 935
QY 38 -----DKTHTCTPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 91
Db 936 ETESGGDKTHTCTPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 995
QY 92 FNMVYDGVVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKKVSNAKALPAPIEK 151
Db 996 FNMVYDGVVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKKVSNAKALPAPIEK 1055
QY 152 TISKAKGQPREPOVYITLPPSRDELTKNQVSLTCLVKGFYPSDIAVEHESNGQPENNYKTT 211
Db 1056 TISKAKGQPREPOVYITLPPSRDELTKNQVSLTCLVKGFYPSDIAVEHESNGQPENNYKTT 1115
QY 212 PPVLDSGGSFFLYSKLTVDSKRWQGNVFCVSMHEALHNHYTKLSLSPGK 264
Db 1116 PPVLDSGGSFFLYSKLTVDSKRWQGNVFCVSMHEALHNHYTKLSLSPGK 1168

RESULT 12
AAG63640
ID AAG63640 standard; Protein; 475 AA.
XX
AC AAG63640;
XX

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:26 ; Search time 14.7529 seconds
(without alignments)
526.515 Million cell updates/sec

Title: 7LINK7LINK2

Perfect score: 1430

Sequence: 1 DMLKAFYDKVAEKLKAEFDW.....MHEALHNHYTKLSLSLSPCK 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1261	88.2	476	3	US-08-487-550-12
2	1260	88.1	449	1	US-08-458-516-13
3	1259	88.0	664	3	US-08-957-063-16
4	1259	88.0	664	3	US-08-957-063-18
5	1259	88.0	664	4	US-09-487-685-16
6	1259	88.0	664	4	US-09-487-685-18
7	1259	88.0	664	4	US-08-802-805D-16
8	1259	88.0	664	4	US-08-802-805D-18
9	1258	88.0	476	2	US-08-378-939-10
10	1257	87.9	437	5	PCT-US96-10043-11
11	1257	87.9	454	2	US-07-934-373C-22
12	1257	87.9	454	3	US-08-437-642B-22
13	1257	87.9	454	4	US-08-146-206C-22
14	1257	87.9	454	5	PCT-US93-07832-22
15	1256	87.8	347	1	US-07-940-861-43
16	1256	87.8	347	1	US-08-459-512-43
17	1256	87.8	347	2	US-08-459-657-43
18	1256	87.8	347	2	US-08-460-132-43
19	1256	87.8	347	4	US-08-466-465-8
20	1256	87.8	347	5	PCT-US92-02050-43
21	1255	87.8	452	3	US-09-027-449-71
22	1255	87.8	452	4	US-09-026-985-71
23	1255	87.8	452	4	US-09-121-952A-71
24	1255	87.8	452	4	US-09-234-340A-71
25	1254	87.7	459	1	US-08-157-101A-7
26	1252	87.6	424	5	PCT-US95-03865-12
27	1252	87.6	424	5	PCT-US95-03866-14

28	1252	87.6	482	4	US-09-189-129-2	Sequence 2, Appli
29	1251	87.5	453	3	US-08-466-151-8	Sequence 8, Appli
30	1251	87.5	453	4	US-08-466-163B-8	Sequence 8, Appli
31	1251	87.5	453	4	US-09-301-593-18	Sequence 18, Appli
32	1251	87.5	467	4	US-09-049-672A-8	Sequence 8, Appli
33	1251	87.5	472	4	US-09-301-593-30	Sequence 30, Appli
34	1251	87.5	911	2	US-08-484-438-10	Sequence 10, Appli
35	1250	87.4	446	3	US-08-397-411-7	Sequence 7, Appli
36	1250	87.4	450	4	US-09-532-856-5	Sequence 5, Appli
37	1250	87.4	476	3	US-08-487-550-4	Sequence 4, Appli
38	1250	87.4	977	4	US-09-590-656-1	Sequence 1, Appli
39	1249	87.3	478	3	US-08-487-550-8	Sequence 8, Appli
40	1248	87.3	388	4	US-09-131-247-16	Sequence 16, Appli
41	1248	87.3	442	5	PCT-US96-10043-9	Sequence 9, Appli
42	1248	87.3	451	2	US-08-887-352B-14	Sequence 14, Appli
43	1248	87.3	451	2	US-08-887-352B-16	Sequence 16, Appli
44	1248	87.3	451	3	US-08-466-151-65	Sequence 65, Appli
45	1248	87.3	451	4	US-09-109-207C-14	Sequence 14, Appli

ALIGNMENTS

RESULT 1
US-08-487-550-12
; Sequence 12, Application US/08487550
; Patent No. 6113898
; GENERAL INFORMATION:
; APPLICANT: Anderson, Darrell R.
; TITLE OF INVENTION: "MONKEY MONOCLONAL ANTIBODIES SPECIFIC
; TO HUMAN B7.1 AND/OR B7.2 PRIMATIZED FORMS THEREOF,
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING, AND USE THEREOF AS
; TITLE OF INVENTION: IMMUNOSUPPRESSANTS"
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: 699 Prince Street
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,550
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 012712-131
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6620
; TELEFAX: 703-836-2021
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 476 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-487-550-12

Query Match 88.2%; Score 1261; DB 3; Length 476;
Best Local Similarity 53.0%; Pred. No. 1.3e-41;
Matches 236; Conservative 7; Mismatches 13; Indels 189; Gaps 2;
QY 9 KVAEKLK-----EAFDWLKAFYD----- 26
| : : : : :
Db 32 KPSETLSLCAVSGSGISGGYWGWIWIRPPKGLGIEWGFSYSSGNTYNNPSLKSQVTIS 91

QY 27 ----- 26
Db 92 TDTSKNQFSLKLSMTAADTAVYICVRDRLESVVGMYNNWFDVWGPGLVTVSSASTKG 151
QY 27 ----- 26
Db 152 PSVFPLAPSSKSTSGTAAALGCLVKDYFPEPTVSWNSGALTSGVHTFPAVLQSSGLYSL 211
QY 27 -----
Db 212 SSVVTVPSSSLGTQTYICNVNHNKPSNTKVDKKAEPKSCDKTHTCPPCPAPPELLGGPSVEL 59
QY 60 FPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNKTKPREQYNSTYRV 119
Db 272 FPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNKTKPREQYNSTYRV 331
QY 120 VSVLTVLHODWLNKGYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQ 179
Db 332 VSVLTVLHODWLNKGYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQ 391
QY 180 VSLTCLVKGFYPSDIAVEHESGQPENNYKTTTPPVLDSDGSFELYSKLTVDKSRWQGNV 239
Db 392 VSLTCLVKGFYPSDIAVEHESGQPENNYKTTTPPVLDSDGSFELYSKLTVDKSRWQGNV 451
QY 240 FSCVMHEALHNHYTKQSLSPGK 264
Db 452 FSCVMHEALHNHYTKQSLSPGK 476

RESULT 2

US-08-458-516-13
; Sequence 13, Application US/08458516
; Patent No. 5777085
; GENERAL INFORMATION:
; APPLICANT: Co, Man Sung
; APPLICANT: TSO, J. Yun
; TITLE OF INVENTION: Humanized Antibodies Reactive with
; TITLE OF INVENTION: GPIIB/IIIA
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: William M. Smith
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,516
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/059,159
; FILING DATE: 03-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11823-37-3
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 449 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-458-516-13

Query Match 88.1%; Score 1260; DB 1; Length 449;
Best Local Similarity 55.9%; Pred. No. 1.3e-41;
Matches 236; Conservative 10; Mismatches 14; Indels 162; Gaps 2;
QY 5 AFYDKVAEKLKEA-----FDWLKAFY----- 25
Db 28 AFTNVLIEWRQAPGQGLEWIGVIYPGSGGTNYNEKFKGRVTLTVDESTNTAYMELSSLR 87
QY 26 ----- 25
Db 88 SEDTAVYFCARRDGNVGNFAYWQGTGLTVTVSSASTKGPSVFFPLAPSSKSTSGTAAALGCL 147
QY 26 ----- 25
Db 148 VKDYFPEPTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVTVPSSSLGTQTYICNVNHNK 207
QY 26 ---DKVAEKLKEAFMDKTHTCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVD 82
Db 208 PSNTKVDKKVEPKSCDKTHTCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVD 267
QY 83 VSHEDPEVKFNWYVDGVEVHNKTKPREQYNSTYRVVSVLTVLHODWLNKGYKCKVSN 142
Db 268 VSHEDPEVKFNWYVDGVEVHNKTKPREQYNSTYRVVSVLTVLHODWLNKGYKCKVSN 327
QY 143 KALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEHESG 202
Db 328 KALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEHESG 387
QY 203 OPENNYKTTTPPVLDSDGSFELYSKLTVDKSRWQGNVSCSVNHEALHNHYTKQSLSP 262
Db 388 OPENNYKTTTPPVLDSDGSFELYSKLTVDKSRWQGNVSCSVNHEALHNHYTKQSLSP 447
QY 263 GK 264
Db 448 GK 449
RESULT 3
US-08-957-063-16
; Sequence 16, Application US/08957063
; Patent No. 6025157
; GENERAL INFORMATION:
; APPLICANT: Robert D. Klein, Arnon Rosenthal, Mary A. Hynes
; TITLE OF INVENTION: Neurturin Receptor
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/957,063
; FILING DATE: 24-Oct-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/871
; FILING DATE: 9-Jun-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 913
; FILING DATE: 18-Feb-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, PhD., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P1086P2
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-8674
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 664 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-957-063-16

Query Match 88.0%; Score 1259; DB 3; Length 664;
Best Local Similarity 51.4%; Pred. No. 3.6e-41;
Matches 237; Conservative 11; Mismatches 14; Indels 199; Gaps 4;

```
QY 3 LKAFYDKV-----AEKLKE-----16
|:|:|:|
DB 204 LROFFDRVPSEYTYRMLFSCDQACAERRRQTILPSCSYEDKEKPNCLDLRGVCRDHL 263
|:|:|:|
QY 17 -----AFDWLKAFYD-----26
DB 264 CRSRLADFHANCRASTYVITSCPADNYQACLSYAGMIGDMTPNYVDSSPTGIVVSPWC 323
|:|:|:|
QY 27 -----KVAEKLKEAF-----36
DB 324 SCRGSGNMECEKEFLRDFTEPCPLRNAIQAFNGTGVNVPKGPSQATQAPRVEKTPS 383
|:|:|:|
QY 37 -----MDKTHTC 43
DB 384 LPDDLSDSTSLGTSVITTCSTVQOGLKANNKSKLSMCFTELTTNIIPGRDPVDKTHTC 443
|:|:|:|
QY 44 PPCAPPELLGGSPVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHN 103
|:|:|:|
DB 444 PPCAPPELLGGSPVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHN 503
|:|:|:|
QY 104 AKTKPREEQNSTYRVVSVLTVLHQLDNLNGKEYCKVSKNKPALPAIEKTSKAGQPREP 163
|:|:|:|
DB 504 AKTKPREEQNSTYRVVSVLTVLHQLDNLNGKEYCKVSKNKPALPAIEKTSKAGQPREP 563
|:|:|:|
QY 164 QVYTLPPSRDELTKNOVSLTCLVKGYFSPDIAEWESNGQPENNYKTPPVLDSDGSFFL 223
|:|:|:|
DB 564 QVYTLPPSRDELTKNOVSLTCLVKGYFSPDIAEWESNGQPENNYKTPPVLDSDGSFFL 623
|:|:|:|
QY 224 YSKLTVDKSRWQOQNVFSCSVMEALHNHYTQKSLSLSPGK 264
|:|:|:|
DB 624 YSKLTVDKSRWQOQNVFSCSVMEALHNHYTQKSLSLSPGK 664
|:|:|:|
```

RESULT 4
US-08-957-063-18

Sequence 18, Application US/08957063
Patent No. 6025157
GENERAL INFORMATION:
APPLICANT: Robert D. Klein, Arnon Rosenthal, Mary A. Hynes
TITLE OF INVENTION: Neurturin Receptor
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/957,063
FILING DATE: 24-Oct-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/871
FILING DATE: 9-Jun-1997

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 913
FILING DATE: 18-Feb-1997
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, PhD., Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: P1086P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-8674
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 664 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-957-063-18

Query Match 88.0%; Score 1259; DB 3; Length 664;
Best Local Similarity 51.4%; Pred. No. 3.6e-41;
Matches 237; Conservative 11; Mismatches 14; Indels 199; Gaps 4;

```
QY 3 LKAFYDKV-----AEKLKE-----16
|:|:|:|
DB 204 LROFFDRVPSEYTYRMLFSCDQACAERRRQTILPSCSYEDKEKPNCLDLRSLCRTDHL 263
|:|:|:|
QY 17 -----AFDWLKAFYD-----26
DB 264 CRSRLADFHANCRASTYVITSCPADNYQACLSYAGMIGDMTPNYVDSSPTGIVVSPWC 323
|:|:|:|
QY 27 -----KVAEKLKEAF-----36
DB 324 SCRGSGNMECEKEFLRDFTEPCPLRNAIQAFNGTGVNVPKGPSQATQAPRVEKTPS 383
|:|:|:|
QY 37 -----MDKTHTC 43
DB 384 LPDDLSDSTSLGTSVITTCSTVQOGLKANNKSKLSMCFTELTTNIIPGRDPVDKTHTC 443
|:|:|:|
QY 44 PPCAPPELLGGSPVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHN 103
|:|:|:|
DB 444 PPCAPPELLGGSPVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHN 503
|:|:|:|
QY 104 AKTKPREEQNSTYRVVSVLTVLHQLDNLNGKEYCKVSKNKPALPAIEKTSKAGQPREP 163
|:|:|:|
DB 504 AKTKPREEQNSTYRVVSVLTVLHQLDNLNGKEYCKVSKNKPALPAIEKTSKAGQPREP 563
|:|:|:|
QY 164 QVYTLPPSRDELTKNOVSLTCLVKGYFSPDIAEWESNGQPENNYKTPPVLDSDGSFFL 223
|:|:|:|
DB 564 QVYTLPPSRDELTKNOVSLTCLVKGYFSPDIAEWESNGQPENNYKTPPVLDSDGSFFL 623
|:|:|:|
QY 224 YSKLTVDKSRWQOQNVFSCSVMEALHNHYTQKSLSLSPGK 264
|:|:|:|
DB 624 YSKLTVDKSRWQOQNVFSCSVMEALHNHYTQKSLSLSPGK 664
|:|:|:|
```

RESULT 5

US-09-487-685-16
Sequence 16, Application US/09487685
Patent No. 6342348
GENERAL INFORMATION:
APPLICANT: Robert D. Klein, Arnon Rosenthal, Mary A. Hynes
TITLE OF INVENTION: Neurturin Receptor
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS


```

; Patent No. 6372453
; GENERAL INFORMATION:
; APPLICANT: Robert D. Klein
; TITLE OF INVENTION: Neurturin Receptor
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/802,805D
; FILING DATE: 18-Feb-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, Ph.D., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P1086
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8674
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 664 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-802-805D-16

```

```

Query Match      88.0%; Score 1259; DB 4; Length 664;
Best Local Similarity 51.4%; Pred. No. 3.6e-41;
Matches 237; Conservative 11; Mismatches 14; Indels 199; Gaps 4;

QY  3 LKAFYDKV-----AEKLE----- 16
   |::|::|
Db  204 LRQFFDRVPSEYTYRMLFCSCQDQACAEARRRQITLPSCSYEDKEKPNCLDLRGVCRTDHL 263
   |::|::|
QY  17 -----AFDWLKAFYD----- 26
   |::|
Db  264 CRSRLADFHANCRASTYRTITSCPADNYQACLSYAGMIGFDMTPNYVDSSPTGIVVSPWC 323
   |::|
QY  27 -----KVAEKLKEAF----- 36
   |::|
Db  324 SCRGSGNMEEECEKFLRDFTEPNCRLNAIQAFNGTGVNVPKGPSFQATAPRVEKTPS 383
   |::|
QY  37 -----MDKTHTC 43
   |::|
Db  384 LPDDLSDSTSLGTSVITTCSTQEOGLKANNKSELSMCFTELTTNIPGRDPVDKTHTC 443
   |::|
QY  44 PPCAPPELLGGSPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHN 103
   |::|
Db  444 PPCAPPELLGGSPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHN 503
   |::|
QY  104 AKTKPREEQNSTYRYVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAKGQPREP 163
   |::|
Db  504 AKTKPREEQNSTYRYVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAKGQPREP 563
   |::|
QY  164 QVYTLPPSRDELTKNQVSLTCLVKGYFSPDIAVEWESNGQPENNYKTPPPVLDSDGSFFL 223
   |::|
Db  564 QVYTLPPSRDEMTKNQVSLTCLVKGYFSPDIAVEWESNGQPENNYKTPPPVLDSDGSFFL 623
   |::|
QY  224 YSKLTVDKSRWQOGNPFSCSVMEALHNYHTQKSLSPCK 264
   |::|
Db  624 YSKLTVDKSRWQOGNPFSCSVMEALHNYHTQKSLSPCK 664

```

RESULT 8

```

US-08-802-805D-18
; Sequence 18, Application US/08802805D
; Patent No. 6372453
; GENERAL INFORMATION:
; APPLICANT: Robert D. Klein
; TITLE OF INVENTION: Neurturin Receptor
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/802,805D
; FILING DATE: 18-Feb-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, Ph.D., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P1086
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8674
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 664 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-802-805D-18

```

```

Query Match      88.0%; Score 1259; DB 4; Length 664;
Best Local Similarity 51.4%; Pred. No. 3.6e-41;
Matches 237; Conservative 11; Mismatches 14; Indels 199; Gaps 4;

QY  3 LKAFYDKV-----AEKLE----- 16
   |::|::|
Db  204 LRQFFDRVPSEYTYRMLFCSCQDQACAEARRRQITLPSCSYEDKEKPNCLDLRLSCLRTDHL 263
   |::|::|
QY  17 -----AFDWLKAFYD----- 26
   |::|
Db  264 CRSRLADFHANCRASTYRTITSCPADNYQACLSYAGMIGFDMTPNYVDSSPTGIVVSPWC 323
   |::|
QY  27 -----KVAEKLKEAF----- 36
   |::|
Db  324 SCRGSGNMEEECEKFLRDFTEPNCRLNAIQAFNGTGVNVPKGPSLPAQAPRVEKTPS 383
   |::|
QY  37 -----MDKTHTC 43
   |::|
Db  384 LPDDLSDSTSLGTSVITTCSTQEOGLKANNKSELSMCFTELTTNIPGRDPVDKTHTC 443
   |::|
QY  44 PPCAPPELLGGSPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHN 103
   |::|
Db  444 PPCAPPELLGGSPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHN 503
   |::|
QY  104 AKTKPREEQNSTYRYVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAKGQPREP 163
   |::|
Db  504 AKTKPREEQNSTYRYVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAKGQPREP 563
   |::|
QY  164 QVYTLPPSRDELTKNQVSLTCLVKGYFSPDIAVEWESNGQPENNYKTPPPVLDSDGSFFL 223
   |::|
Db  564 QVYTLPPSRDEMTKNQVSLTCLVKGYFSPDIAVEWESNGQPENNYKTPPPVLDSDGSFFL 623
   |::|
QY  224 YSKLTVDKSRWQOGNPFSCSVMEALHNYHTQKSLSPCK 264
   |::|
Db  624 YSKLTVDKSRWQOGNPFSCSVMEALHNYHTQKSLSPCK 664

```

RESULT 9
US-08-378-939-10
; Sequence 10, Application US/08378939
; Patent No. 5876961
; GENERAL INFORMATION:
; APPLICANT: CROWE, JAMES SCOTT
; APPLICANT: LEWIS, ALAN PETER
; TITLE OF INVENTION: PRODUCTION OF ANTIBODIES
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROTHWELL, FIGG, ERNST & KURZ
; STREET: 555 THIRTEENTH ST. N.W.
; CITY: WASHINGTON
; STATE: D. C.
; COUNTRY: U.S.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/378.939
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/952640
; FILING DATE: 01-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: ERNST, BARBARA G
; REGISTRATION NUMBER: 30,377
; REFERENCE/DOCKET NUMBER: 1808-118
; TELEPHONE: (202) 783-6040
; TELEFAX: (202) 783-6031
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 476 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-378-939-10

Query Match 88.0%; Score 1258; DB 2; Length 476;
Best Local Similarity 57.8%; Pred. No. 1.8e-41;
Matches 238; Conservative 9; Mismatches 17; Indels 148; Gaps 4;
QY 1 DWL-----KAFYDKVAEKLKEA- 17
Db 65 EWMGIIPLFGTPTYSQNFQGRVITADKSTSTAHMELTSLRSEDYAVVYCATDRYQAN 124
QY 18 FD-----WL----- 21
Db 125 FDRARVGFDPWGQGLTVTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPTV 184
QY 22 -----KAFYDKVAEKL 32
Db 185 VSWNSGALTSGVHFFPAVLQSSGLYSLSVTVVPSSSLGTTQTYICNVNHNKPSNTKVDKKV 244
QY 33 KEAFMDKTHTCPAPPELLGGPSVFLFPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKF 92
Db 245 EPKSCDKTHTCPAPPELLGGPSVFLFPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKF 304
QY 93 NWYDGVGVHNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 152
Db 305 NWYDGVGVHNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKT 364
QY 153 ISKAGQPREQVYTLPPSRDELTKNOVSLTCLVKGFPYPSDIAVWESNGQPENNYKTP 212
Db 365 ISKAGQPREQVYTLPPSRDELTKNOVSLTCLVKGFPYPSDIAVWESNGQPENNYKTP 424
QY 213 PVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 264

Db 425 PVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 476
RESULT 10
PCT-US96-10043-11
; Sequence 11, Application PC/TUS9610043
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation
; TITLE OF INVENTION: P-SELECTIN LIGANDS AND RELATED MOLECULES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02210-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10043
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/000,213
; FILING DATE: 14-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lech, Karen F.
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 00786/284001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US96-10043-11

Query Match 87.9%; Score 1257; DB 5; Length 437;
Best Local Similarity 79.5%; Pred. No. 1.6e-41;
Matches 241; Conservative 6; Mismatches 13; Indels 43; Gaps 6;
QY 1 DW-LKAFYDK---VAEKLKEAF-----DWLKAFYDKVAEKLKEAF----- 36
Db 139 NWGLSVADKPTTKQELGEFEYALDCLRIKPSDVVYTDWKK---DK-CEPLEKHOKER 194
QY 37 -----MDKTHTCPAPPELLGGPSVFLFPPPKKDTLMISRTPEVTCVV 81
Db 195 KOEGESDPEGEKSCDKTHTCPAPPELLGGPSVFLFPPPKKDTLMISRTPEVTCVV 254
QY 82 DVSHEDPEVKFNWYVDGVEVHNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVS 141
Db 255 DVSHEDPEVKFNWYVDGVEVHNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVS 314
QY 142 NKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESN 201
Db 315 NKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESN 374
QY 202 GQPNNTKTPPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLS 261
Db 375 GQPNNTKTPPVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLS 434

```
QY 262 PGK 264
Db 435 PGK 437

RESULT 11
US-07-934-373C-22
; Sequence 22, Application US/07934373C
; Patent No. 5821337
; GENERAL INFORMATION:
; APPLICANT: Paul J. Carter
; APPLICANT: Leonard G. Presta
; TITLE OF INVENTION: Immunoglobulin Variants
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/934,373C
; FILING DATE: 21-Aug-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05126
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0709P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-07-934-373C-22

Query Match 87.9%; Score 1257; DB 2; Length 454;
Best Local Similarity 54.9%; Pred. No. 1.7e-41;
Matches 234; Conservative 11; Mismatches 11; Indels 170; Gaps 3;

QY 2 WLKAFYDKVAEKLKEAFDWLAKF----- 24
Db 36 WMKQSHGK-----SLEWIGGFNPKNGSSHNORFMDKATLAYDKSTAYMELRSLTS 88

QY 25 -----YD----- 26
Db 89 EDGIYYCARWRGLNYGDFVRYEDVWGAGTTVTVSSASTKGPSVFLPAPSSKSTSGTAA 148

QY 27 ----- 26
Db 149 LGCLVKDYFPEPTVTSNNSGALTSGVHTFPAVLQSSGLYSLSSVTVTPSSSLGTQTVICN 208

QY 27 -----KVAELKEAFMDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTC 78
Db 209 VNHKPSNTKVDKKVEPKSCDKTHCTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTC 268

QY 79 VVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYKC 138
Db 269 VVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYKC 328

QY 139 KVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEW 198
Db 329 KVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDEMTKKNQVSLTCLVKGFYPSDIAVEW 388

QY 199 ESNQGPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQOGNVFSCSVMEALHNNHYTKSL 258
Db 389 ESNQGPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQOGNVFSCSVMEALHNNHYTKSL 448

QY 259 SLSPGK 264
Db 449 SLSPGK 454

RESULT 12
US-08-437-642B-22
; Sequence 22, Application US/08437642B
; Patent No. 6054297
; GENERAL INFORMATION:
; APPLICANT: Paul J. Carter
; APPLICANT: Leonard G. Presta
; TITLE OF INVENTION: Immunoglobulin Variants
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/437,642B
; FILING DATE: 09-May-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934373
; FILING DATE: 21-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/146206
; FILING DATE: 17-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05126
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0709P2C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-437-642B-22

Query Match 87.9%; Score 1257; DB 3; Length 454;
Best Local Similarity 54.9%; Pred. No. 1.7e-41;
Matches 234; Conservative 11; Mismatches 11; Indels 170; Gaps 3;

QY 2 WLKAFYDKVAEKLKEAFDWLAKF----- 24
Db 36 WMKQSHGK-----SLEWIGGFNPKNGSSHNORFMDKATLAYDKSTAYMELRSLTS 88
```

QY 25 -----YD----- 26
Db 89 EDGIIYCARWGLNYGDFVYDFVWGAGTTVTVSSASTKGPSVFLPAPSSKSTSGGTAA 148
QY 27 ----- 26
Db 149 LGCLVKDYFPEPVTVSWNSGALTSVHTTTPAVLQSSGLYSLSSVTVTPSSSLGTQTYICN 208
QY 27 -----KVAEKLKEAFMDKTHCPCPAPELLGGPSVFLFPKPKDFTLMISRTPPEVTC 78
Db 209 VNHKPSNTKVDKVEPKSCDKTHCPCPAPELLGGPSVFLFPKPKDFTLMISRTPPEVTC 268
QY 79 VYDVSHEDPEVKFNMYVDGVEVHNNAKTPREQYNSTYRVSVLTVLHODWMLNGKEYC 138
Db 269 VYDVSHEDPEVKFNMYVDGVEVHNNAKTPREQYNSTYRVSVLTVLHODWMLNGKEYC 328
QY 139 KVSNAKALPAPIETISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEW 198
Db 329 KVSNAKALPAPIETISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEW 388
QY 199 EMSGOPENNYKTPPPVLDSDGSFFLYSKLTVDKSRWQOGNVFSCSVMHHEALHNNHYTQKSL 258
Db 389 EMSGOPENNYKTPPPVLDSDGSFFLYSKLTVDKSRWQOGNVFSCSVMHHEALHNNHYTQKSL 448
QY 259 SLSPGK 264
Db 449 SLSPGK 454

RESULT 13

US-08-146-206C-22
; Sequence 22, Application US/08146206C
; Patent No. 6407213
; GENERAL INFORMATION:
; APPLICANT: Carter, Paul J.
; APPLICANT: Presta, Leonard G.
; TITLE OF INVENTION: Method for Making Humanized Antibodies
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,206C
; FILING DATE: 17-No. 6407213-1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0709P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-146-206C-22

Query Match 87.9%; Score 1257; DB 4; Length 454;
Best Local Similarity 54.9%; Pred. No. 1.7e-41;

Matches 234; Conservative 11; Mismatches 11; Indels 170; Gaps 3;
QY 2 WLKAFDYKVAEKLKEAFMDLKAF----- 24
Db 36 WVKQSHGK-----SLEWIGGFNPKNGGSSHNQRFMDKATLAVDKSTSTAYMELRSLTS 88
QY 25 -----YD----- 26
Db 89 EDGIIYCARWGLNYGDFVYDFVWGAGTTVTVSSASTKGPSVFLPAPSSKSTSGGTAA 148
QY 27 ----- 26
Db 149 LGCLVKDYFPEPVTVSWNSGALTSVHTTTPAVLQSSGLYSLSSVTVTPSSSLGTQTYICN 208
QY 27 -----KVAEKLKEAFMDKTHCPCPAPELLGGPSVFLFPKPKDFTLMISRTPPEVTC 78
Db 209 VNHKPSNTKVDKVEPKSCDKTHCPCPAPELLGGPSVFLFPKPKDFTLMISRTPPEVTC 268
QY 79 VYDVSHEDPEVKFNMYVDGVEVHNNAKTPREQYNSTYRVSVLTVLHODWMLNGKEYC 138
Db 269 VYDVSHEDPEVKFNMYVDGVEVHNNAKTPREQYNSTYRVSVLTVLHODWMLNGKEYC 328
QY 139 KVSNAKALPAPIETISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEW 198
Db 329 KVSNAKALPAPIETISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEW 388
QY 199 EMSGOPENNYKTPPPVLDSDGSFFLYSKLTVDKSRWQOGNVFSCSVMHHEALHNNHYTQKSL 258
Db 389 EMSGOPENNYKTPPPVLDSDGSFFLYSKLTVDKSRWQOGNVFSCSVMHHEALHNNHYTQKSL 448
QY 259 SLSPGK 264
Db 449 SLSPGK 454

RESULT 14

PCT-US93-07832-22
; Sequence 22, Application PC/TUS9307832
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: Immunoglobulin Variants
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07832
; FILING DATE: 19930820
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05126
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934373
; FILING DATE: 21-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 709P2PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:

TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 454 amino acids
TYPE: amino acid
TOPOLOGY: linear
PCT-US93-07832-22

Query Match 87.9%; Score 1257; DB 5; Length 454;
Best Local Similarity 54.9%; Pred. No. 1.7e-41;
Matches 234; Conservative 11; Mismatches 11; Indels 170; Gaps 3;
QY 2 WLKAFYDKVAEKLKEAFDMLKAF----- 24
Db 36 WMKQSHGK-----SLEWIGGPNKGGSHNQRFMDKATLAVDKSTSTAYMELRSLTS 88
QY 25 -----YD----- 26
Db 89 EDGSIYYCARWGLNYGFDVRYEDVWAGTTVTVSSASTKGPSVFLPAPSSKSTSGGTAA 148
QY 27 ----- 26
Db 149 LGLVLDYPPPTVTVSNWNSGALTSVHTPPAVLQSSGLYSLSVTVTPSSSLGTQTYICN 208
QY 27 -----KVAEKLKEAFMDKHTCCPCPAPELLGGPSVFLFPKPKDITLMISRTPEVTC 78
Db 209 VNHKPSNTKVDKVEPKSCDKHTCCPCPAPELLGGPSVFLFPKPKDITLMISRTPEVTC 268
QY 79 VVDVSHEDPEVFNMYVDGVEVHNNAKTPREQYNSTYRVSVLTVLHODWLNKGEYKC 138
Db 269 VVDVSHEDPEVFNMYVDGVEVHNNAKTPREQYNSTYRVSVLTVLHODWLNKGEYKC 328
QY 139 KVNKALPAPIEKTISKAGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAYEW 198
Db 329 KVNKALPAPIEKTISKAGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAYEW 388
QY 199 ESNQPPENNYKTPPPVLDSDGSPFLYSLKLTVDKSRWQOGNVFSCSVNHEALHNNHYTKQSL 258
Db 389 ESNQPPENNYKTPPPVLDSDGSPFLYSLKLTVDKSRWQOGNVFSCSVNHEALHNNHYTKQSL 448
QY 259 SLSPGK 264
Db 449 SLSPGK 454

RESULT 15
US-07-940-861-43
Sequence 43, Application US/07940861
Patent No. 5547853
GENERAL INFORMATION:
APPLICANT: BIOGEN, INC.
APPLICANT: WALLNER, Barbara P.
APPLICANT: MILLER, Glenn T.
APPLICANT: ROSA, Margaret D.
TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Neave
STREET: 875 Third Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10022-6250
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/940,861
FILING DATE: 21-OCT-1992

CLASSIFICATION: 514
PRIOR APPLICATION DATA: PCT/US92/02050
FILING DATE: 12-MAR-1992
PRIOR APPLICATION DATA: US 07/667,971
FILING DATE: 12-MAR-1991
PRIOR APPLICATION DATA: US 07/770,967
FILING DATE: 07-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: HALEY, James F., Jr.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B151CIP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)715-0600
TELEFAX: (212)715-0673
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 347 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-940-861-43
Query Match 87.8%; Score 1256; DB 1; Length 347;
Best Local Similarity 78.5%; Pred. No. 9.8e-42;
Matches 238; Conservative 3; Mismatches 11; Indels 51; Gaps 3;
QY 2 WLKAFYDKVAEKLKEAF-----DWL 21
Db 56 WKQK-KDKVAELENSEFRAPSSFKNRVYLDTVSGSLTYINLTSSDEDEYEMESPNTDTM 114
QY 22 KAFYDKVAEKLKEAFMDKHTCCPCPAPELLGGPSVFLFPKPKDITLMISRTPEVTCVV 81
Db 115 KFF-----LYVDKHTCCPCPAPELLGGPSVFLFPKPKDITLMISRTPEVTCVV 164
QY 82 DVSHEDPEVKFNMYVDGVEVHNNAKTPREQYNSTYRVSVLTVLHODWLNKGEYKC 141
Db 165 DVSHEDPEVKFNMYVDGVEVHNNAKTPREQYNSTYRVSVLTVLHODWLNKGEYKC 224
QY 142 NKALPAPIEKTISKAGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAYEWESN 201
Db 225 NKALPAPIEKTISKAGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAYEWESN 284
QY 202 GOPENNYKTPPPVLDSDGSPFLYSLKLTVDKSRWQOGNVFSCSVNHEALHNNHYTKQSL 261
Db 285 GOPENNYKTPPPVLDSDGSPFLYSLKLTVDKSRWQOGNVFSCSVNHEALHNNHYTKQSL 344
QY 262 PGK 264
Db 345 PGK 347

Search completed: April 21, 2003, 10:50:25
Job time : 16.7529 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:43:21 ; Search time 12.1647 Seconds
(without alignments)
1640.982 Million cell updates/sec

Title: 7LINK7LINK2
Perfect score: 1430
Sequence: 1 DWLAFYDKVAEKLKEAFDW.....MHEALHNYHTOKLSLSLSPGK 264

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.0

Searched: 288829 seqs, 75613885 residues
Total number of hits satisfying chosen parameters: 288829

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/ECT_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/ECTUS_PUBCOMB pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1268	88.7	594	10 US-09-815-108-22	Sequence 22, Appli
2	1268	88.7	951	9 US-09-935-868-9	Sequence 9, Appli
3	1268	88.7	1158	9 US-09-935-868-26	Sequence 26, Appli
4	1268	88.7	1168	9 US-09-935-868-24	Sequence 24, Appli
5	1263	88.3	567	10 US-09-825-561A-16	Sequence 16, Appli
6	1263	88.3	859	9 US-09-935-868-7	Sequence 7, Appli
7	1261	88.2	476	9 US-10-124-905-12	Sequence 12, Appli
8	1261	88.2	476	9 US-09-948-429B-12	Sequence 12, Appli
9	1260	88.1	399	9 US-09-832-659-2	Sequence 2, Appli
10	1259	88.0	547	10 US-09-746-359A-54	Sequence 54, Appli
11	1259	88.0	571	10 US-09-746-359A-53	Sequence 53, Appli
12	1258	88.0	475	10 US-09-740-002-27	Sequence 27, Appli
13	1256	87.8	347	9 US-10-091-236-17	Sequence 17, Appli
14	1256	87.8	347	9 US-10-091-313-7	Sequence 7, Appli
15	1256	87.8	347	9 US-10-091-268-7	Sequence 7, Appli
16	1256	87.8	347	10 US-09-796-033-8	Sequence 8, Appli
17	1256	87.8	347	10 US-09-730-465-8	Sequence 8, Appli
18	1256	87.8	776	9 US-09-935-868-36	Sequence 36, Appli
19	1256	87.8	776	9 US-09-935-868-40	Sequence 40, Appli

20	1256	87.8	776	9 US-09-935-868-44	Sequence 44, Appli
21	1256	87.8	780	9 US-09-935-868-34	Sequence 34, Appli
22	1256	87.8	780	9 US-09-935-868-38	Sequence 38, Appli
23	1256	87.8	780	9 US-09-935-868-42	Sequence 42, Appli
24	1256	87.8	910	9 US-09-935-868-28	Sequence 28, Appli
25	1256	87.8	423	9 US-09-832-659-44	Sequence 44, Appli
26	1255	87.8	452	9 US-09-726-258-71	Sequence 71, Appli
27	1255	87.8	475	10 US-09-740-002-25	Sequence 25, Appli
28	1254	87.7	592	9 US-09-935-868-8	Sequence 8, Appli
29	1254	87.7	691	9 US-09-935-868-20	Sequence 20, Appli
30	1254	87.7	694	9 US-09-935-868-22	Sequence 22, Appli
31	1254	87.7	1194	9 US-10-191-029-10	Sequence 10, Appli
32	1253	87.6	418	9 US-09-832-659-42	Sequence 42, Appli
33	1253	87.6	731	10 US-09-825-012-46	Sequence 46, Appli
34	1253	87.6	741	10 US-09-825-012-55	Sequence 55, Appli
35	1252	87.6	450	9 US-09-996-288-228	Sequence 228, App
36	1252	87.6	450	10 US-09-796-848A-47	Sequence 47, Appli
37	1252	87.6	482	10 US-09-824-286-2	Sequence 2, Appli
38	1251	87.5	450	9 US-09-996-288-216	Sequence 216, App
39	1251	87.5	450	9 US-09-996-288-218	Sequence 218, App
40	1251	87.5	450	9 US-09-996-288-250	Sequence 250, App
41	1251	87.5	450	10 US-09-796-848A-43	Sequence 43, Appli
42	1251	87.5	450	10 US-09-796-848A-49	Sequence 49, Appli
43	1251	87.5	451	10 US-09-822-698A-26	Sequence 26, Appli
44	1251	87.5	453	9 US-09-925-179-8	Sequence 8, Appli
45	1251	87.5	453	10 US-09-802-077-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-815-108-22
; Sequence 22, Application US/09815108
; Patent No. US20020009776A1
; GENERAL INFORMATION:
; APPLICANT: Saris, Christiaan M.
; APPLICANT: Sharon, Mu X.
; APPLICANT: Xia, Min
; APPLICANT: Boone, Thomas Charles
; APPLICANT: Covey, Todd
; TITLE OF INVENTION: Fibroblast Growth Factor Receptor-Like Molecules and
; TITLE OF INVENTION: Uses Thereof
; FILE REFERENCE: 99-513-A
; CURRENT APPLICATION NUMBER: US/09/815,108
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/191,379
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: murine FGFR-L
; OTHER INFORMATION: extracellular domain-Fc fusion polypeptide
US-09-815-108-22

Query Match	88.7%	Score 1268;	DB 10;	Length 594;
Best Local Similarity	57.6%;	Pred. No. 3.3e-32;		
Matches 239;	Conservative	9;	Mismatches 15;	Indels 152; Gaps 4;
QY	2	WLK-----		4
		!::		
Db	180	WMKDDQTLTLEASERHKRWTLKLNLPEDSGKTCRVSNKAGAINATYKVDVIQTR	239	
QY	5	-----AFYDKVAEKLKEAFDWLK-----		22
		!::		
Db	240	SKPVLGTGTHPVNTVDFGGTTSFQCKVRSDVKPVQIWLKKRVEYSGRHNSTIDVGQKF	299	
QY	23	-----AFYDK-----		37
		!::		

Db	300	VVLPTGVDWSPDGSYLNKLLISRARODDAGMYICLGANTMGVYSFSAFVLTVLPDPKPPG	359
Qy	38	-----DKTHTCCPPAPPELLGGPSVFLFPKPKDTLMI	89
Db	360	PMMASSVDKTHTCPPCAPPELLGGPSVFLFPKPKDTLMI	419
Qy	90	VKFNWYVDGVEVHNAKTKPREEQYNSTIRYVSVLTVLHQDLWNGKEYCKKVS	149
Db	420	VKFNWYVDGVEVHNAKTKPREEQYNSTIRYVSVLTVLHQDLWNGKEYCKKVS	479
Qy	150	EKTSKAGQPREPOVYTLPPSRDELTKNOVSLTCLVKGPYSDIAVESNGSQP	209
Db	480	EKTSKAGQPREPOVYTLPPSRDELTKNOVSLTCLVKGPYSDIAVESNGSQP	539
Qy	210	TTPLVLSDSGFFLYSLKLTVDKSRWQOGNVFSCVMHEALHNHYTKQSL	264
Db	540	TTPLVLSDSGFFLYSLKLTVDKSRWQOGNVFSCVMHEALHNHYTKQSL	594

```

RESULT 2
US-09-935-868-9
; Sequence 9, Application US/09935868
; Patent No. US20020164690A1
; GENERAL INFORMATION:
; APPLICANT: Regeneron Pharmaceuticals, Inc
; TITLE OF INVENTION: Receptor Based Antagonists, and Methods of Making and Using
; FILE REFERENCE: REG 203D
; CURRENT APPLICATION NUMBER: US/09/935,868
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: PCT/US99/22045
; PRIOR FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 951
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-935-868-9

```

```

Query Match      88.7%; Score 1268; DB 9; Length 951;
Best Local Similarity 38.4%; Pred. No. 9.7e-32;
Matches 246; Conservative 8; Mismatches 10; Indels 376; Gaps 7;

Qy      1 DW----- 2
      ||
Db 312 DWSEASGITYEDRPSKAPSWYKIDPSHTQGYTQVLVWMTLPPFANGKILDYEVTLT 371
      ||
Qy      3 -----LK 4
      ||
Db 372 RWKSHLQNYTVNATKLTVNLNTDTRYLATLTVRNLVSGSDAAVLTI PACDFQATHPVMDLK 431
      ||
Qy      5 AF-----YDK-----VAEK-- 13
      ||
Db 432 AFPKDNMLWVEWTTTPRESVKYKYLWCVLSDKAPCITDWOQDGTGVHRTVLRGNLAESKC 491
      ||
Qy      14 -----LKEA-----FDW----- 20
      ||
Db 492 YLITVPYADGPGSPESIKAYLKAQAPSPKPTVTRTKVKGNKNEAVLEMDQLPVDVQNGFI 551
      ||
Qy      21 -----LKFYD----- 26
      ||
Db 552 RNYTIFYRTIIGNETAVNDSSHTEYTLSSLTSDTLVYMRMAAYTDEGGKDPETFTP 611
      ||
Qy      27 ----- 26
      ||
Db 612 KFAQGEIESGASTKGPVFFPLAPSSKSTSGTAALGCLVKDYFPEPTVTVSNNSGALTSGV 671
      ||
Qy      27 -----KVAEKLKEAFMDKTHTCP 44
      ||
Db 672 HTPPAVLQSSGLYSLSWVTVPPSSSLGTQYICNVNHKPSNTKVDDKVEPKSDKTHTCP 731
      ||
Qy      45 PCPAPELLGGPSVFLFPPKPKDTLMIISRTPEVTCVVYDVDSHEDPEVKFNKYVDGVEYHNA 104
      ||

```

Db	732	PCPAPPELLGGPSVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNA	791
Qy	105	KTKPREQYNSTRYRVSVLTVLHODWLGNGKEYCKKSVKNKALPAPIEKTISKAKGQPREPQ	164
Db	792	KTKPREQYNSTRYRVSVLTVLHODWLGNGKEYCKKSVKNKALPAPIEKTISKAKGQPREPQ	851
Qy	165	VYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLY	224
Db	852	VYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLY	911
Qy	225	SKLTVDKSRWQGNVFSCVMEALHNHYTOKSLSLSPGK	264
Db	912	SKLTVDKSRWQGNVFSCVMEALHNHYTOKSLSLSPGK	951

```

RESULT 3
US-09-935-868-26
; Sequence 26, Application US/09935868
; Patent No. US20020164690A1
; GENERAL INFORMATION:
; APPLICANT: Regeneron Pharmaceuticals, Inc
; TITLE OF INVENTION: Receptor Based Antagonists, and Methods of Making and Using
; FILE REFERENCE: REG 203D
; CURRENT APPLICATION NUMBER: US/09/935.868
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: PCT/US99/22045
; PRIOR FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
; LENGTH: 1158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-935-868-26

```

Query Match	88.7%	Score	1268;	DB	9;	Length	1158;
Best Local Similarity	37.1%	Pred. No.	1.5e-31;				
Matches	242;	Conservative	12;	Mismatches	9;	Indels	390;
Gaps	6;						
Qy	2	WLKA-----FVDKV-----AEKL-----	14				
		:	:				
Db	506	WVEAENALGKVTSDHINDPVPYKYPNPNLSVINSELSILKLTWTNPSIKSVILK	565				
Qy	15	-----KEAFDM-----	20				
		:	:				
Db	566	YNIQYTKDASTWSQIOPEDTASTRSSTFQDLKPFTEYVFRICMKEDGKYWSDWSEE	625				
Qy	21	-----	20				
Db	626	ASGTYEDRPSKAPSFWKIDPSHTQGYRTVOLVWKTLPPPEANGKILDYEVTLTRWKSH	685				
Qy	21	-----LKAF-----	24				
Db	686	LQNYTVNATKLTVNLTDNRYLATLTVRNLVGKSDAAVLTIPACDFAQTHPYMDLKAPPKD	745				
Qy	25	-----	24				
Db	746	NMLVWEVTTPRESVKKYILEWCVLSDKAPCITDQOQEDGVHRTYLRGNLAESKCYLITV	805				
Qy	25	-----YDKVAEKLKEAF-----	37				
		:	:				
Db	806	TPVYADGPGSPESIKAYLKQAPPSKGPVTRTKYGVKNEAVLENDQLPVDVQNGFIRNYTI	865				
Qy	38	-----	37				
Db	866	FYRTIIGNETAVNVDSHTEYTLSSLTSDTLYVMRMAAYTDEGGKDGPEFTFTPKFAQG	925				
Qy	38	-----DKTHTCPCPAPPELLGGPSVLFPPKPKDQDTLMISRTPEVTCVVVDVSHEDPEVK	91				
Db	926	BIESGDKTHTCPCPAPPELLGGPSVLFPPKPKDQDTLMISRTPEVTCVVVDVSHEDPEVK	985				
Qy	92	FNWYVDGVEVHENAKTKPREQYNSYRIVSVSLTVLHODWLNGKEYKCKVSNKALPAPIEK	151				

Db 986 FNNYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 1045
QY 152 TISKAGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 211
Db 1046 TISKAGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 1105
QY 212 PPVLDSGSEFFLYSKLTVDKSRWQOGNVSFSCVMHEALHNNHYTQKLSLSLSPGK 264
Db 1106 PPVLDSGSEFFLYSKLTVDKSRWQOGNVSFSCVMHEALHNNHYTQKLSLSLSPGK 1158

RESULT 4
US-09-935-868-24
; Sequence 24, Application US/09935868
; Patent No. US20020164690A1
; GENERAL INFORMATION:
; APPLICANT: Regeneron Pharmaceuticals, Inc
; TITLE OF INVENTION: Receptor Based Antagonists, and Methods of Making and Using
; FILE REFERENCE: REG 203D
; CURRENT APPLICATION NUMBER: US/09/935,868
; PRIOR FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: PCT/US99/22045
; PRIOR FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24
; LENGTH: 1168
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-935-868-24

Query Match 88.7%; Score 1268; DB 9; Length 1168;
Best Local Similarity 37.1%; Pred. No. 1.6e-31;
Matches 242; Conservative 12; Mismatches 9; Indels 390; Gaps 6;

QY 2 WLKA-----FYDKV-----AEKL----- 14
|::| |::| |::|
Db 516 WVEAENALGKVTSDHINFDPYKVPKPNPHNLSVINSEELSLKLTWTNPSIKSVILK 575
|::| |::| |::|

QY 15 -----KEAFDW----- 20
|::| |::| |::|

Db 576 YNIQRTKDASTWSQIPPEDTASTRSFTVODLKPTEYVFRICMKEDGKGYWSDMSEE 635
QY 21 ----- 20

Db 636 ASGITVEDPSKAPSEWKIDPSHTQGYRTVQLVWKTLPPFEANGKILDYEVTLTRWKSH 695
QY 21 -----LKAF----- 24
|::| |::| |::|

Db 696 LQNTVNAKTLVNTLNDRYLATLTVRNLVGKSDAAVLTIPACDFQATHPVMDLKAPPKD 755
QY 25 ----- 24

Db 756 NMLVWEVTTPRESVKKYLEWCVLSDKAPCITDQOQEDGTVHRTYLRGNLAESKCYLITV 815
QY 25 -----YDKVAEKLKEAFM----- 37
|::| |::| |::|

Db 816 TPVIADPGSPESIKAYLQAPPSKGPVTYRTKVGKNEAVLEWDQLPVDVQNGFIRNYTI 875
QY 38 ----- 37

Db 876 FYRTIIGNETAVNVDSSTHTYTLSSLTSDPLYMVRMAAYTDEGKDGKDPETFTTPKPAQG 935
QY 38 -----DKHTCTPCPAPELLGSPSVFLPPPKPDTLMISRTPEVTCVVVDVSHEDPEVK 91
|::| |::| |::|

Db 936 EIESGGDKTHCTPCPAPELLGSPSVFLPPPKPDTLMISRTPEVTCVVVDVSHEDPEVK 995
QY 92 FNNYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 151
|::| |::| |::|

Db 996 FNNYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEK 1055
QY 152 TISKAGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 211
|::| |::| |::|

Db 1056 TISKAGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 1115
QY 212 PPVLDSGSEFFLYSKLTVDKSRWQOGNVSFSCVMHEALHNNHYTQKLSLSLSPGK 264
|::| |::| |::|

Db 1116 PPVLDSGSEFFLYSKLTVDKSRWQOGNVSFSCVMHEALHNNHYTQKLSLSLSPGK 1168

RESULT 5
US-09-825-561A-16
; Sequence 16, Application US/09825561A
; Patent No. US20020137677A1
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy A.
; APPLICANT: No. US20020137677Alak, Julia E.
; APPLICANT: West, James W.
; APPLICANT: Presnell, Scott R.
; APPLICANT: Holly, Richard D.
; APPLICANT: Nelson, Andrew J.
; TITLE OF INVENTION: SOLUBLE ZALPHALL CYTOKINE RECEPTORS
; FILE REFERENCE: 00-22
; CURRENT APPLICATION NUMBER: US/09/825,561A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/194,731
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/222,121
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: soluble zalphallR/IgGgammaal polypeptide
US-09-825-561A-16

Query Match 88.3%; Score 1263; DB 10; Length 567;
Best Local Similarity 56.8%; Pred. No. 4.2e-32;
Matches 241; Conservative 8; Mismatches 9; Indels 166; Gaps 5;

QY 5 AFY----- 7
|::| |::| |::|

Db 146 AFYMLKGLQYLRNRGDPWAVSPRRKLISVDSRSVLSLLPLEFRKDSYELQVRAGPM 205
QY 8 -----DKV-----AEKLKEAFDW----- 20
|::| |::| |::|

Db 206 PGSSYQGTWSEWSDPVIFQTSBELKEG--WNPHASTKGPSVFPPLAPSSKTSKGTAAIG 263
QY 21 --LKAFY----- 25
|::| |::| |::|

Db 264 CLVKDYFPPEVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVTVPSSSLGTQTYICNVN 323
QY 26 -----DKVAEKLKEAFMDKTHCTPCPAPELLGSPSVFLPPPKPDTLMISRTPEVTCVV 80
|::| |::| |::|

Db 324 HKPSNTKVDKVKPEKSCDKTHTCPCPAPELLGSPSVFLPPPKPDTLMISRTPEVTCVV 383
QY 81 VDVSHEDEPEVKFNNYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKV 140
|::| |::| |::|

Db 384 VDVSHEDEPEVKFNNYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKV 443
QY 141 SNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWES 200
|::| |::| |::|

Db 444 SNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWES 503
QY 201 NGQPENNYKTTTPVLDSGSEFFLYSKLTVDKSRWQOGNVSFSCVMHEALHNNHYTQKLSL 260
|::| |::| |::|

Db 504 NGQPENNYKTTTPVLDSGSEFFLYSKLTVDKSRWQOGNVSFSCVMHEALHNNHYTQKLSL 563
QY 261 SPGK 264
|::| |::| |::|

Db 564 SPGK 567

RESULT 6

```

US-09-935-868-7
; Sequence 7, Application US/09935868
; Patent No. US20020164690A1
; GENERAL INFORMATION:
; APPLICANT: Regeneron Pharmaceuticals, Inc
; TITLE OF INVENTION: Receptor Based Antagonists, and Methods of Making and Using
; FILE REFERENCE: REG 203D
; CURRENT APPLICATION NUMBER: US/09/935,868
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: PCT/US99/22045
; PRIOR FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 859
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-935-868-7

```

Query Match 88.3%; Score 1263; DB 9; Length 859;
Best Local Similarity 44.3%; Pred. NO. 1.1e-31;
Matches 240; Conservative 12; Mismatches 12; Indels 278; Gaps 5;

QY	1	DW-----	2
Db	312	DWSEASGITYEDRPSKAPSFWKIDPSHTQGYRTVQLVWKTLPPPEANGKILDYEVLUT	371
QY	3	-----LK 4	4
Db	372	RWKSHLQNYTVNATKTLTVNLNTDRIYIATITVRNLVGKSDAAVLITIPACDFQATHPVMDUK	431
QY	5	AF-----YDKVAEKLKEAFD-	20
Db	432	AFPKDNMLWVEWTTPRESVKYKYLEWCVLSDRAPCITDQOQDGTVHRTYTLRGNLAESKC	491
QY	21	-----LKA-	37
Db	492	YLITVTPVYADGSPSESIKAYLKQAPPKSGPTVRTKKVGKEAVLEWDQLPVDQNGFI	551
QY	38	-----	37
Db	552	RNTIIFYRIIGNETAVNDSSHTEYTLSSLTSDTLIYWRMAAYTDEGKGDPETFTTTP	611
QY	38	-----DKTHTCPAPPELLGGPSVFLFPKPKDITLMISRTPEVTCVVVD	82
Db	612	KFAQGEISEGPKSCDKTHTCPAPPELLGGPSVFLFPKPKDITLMISRTPEVTCVVVD	671
QY	83	VSHEDDEVKFNWYDGVVEVHNAKTPREBQYNSTYRWVSVLTVLHQDWLNGKEYKCKVSN	142
Db	672	VSHEDDEVKFNWYDGVVEVHNAKTPREBQYNSTYRWVSVLTVLHQDWLNGKEYKCKVSN	731
QY	143	KALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLIVKGFFPSDIAVWESNG	202
Db	732	KALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLIVKGFFPSDIAVWESNG	791
QY	203	QPENNYKTTPPVLDSDGSFELYSKLTVDKSRWQQGNVFCSVMHBAALHNHYTQKSLSLSP	262
Db	792	QPENNYKTTPPVLDSDGSFELYSKLTVDKSRWQQGNVFCSVMHBAALHNHYTQKSLSLSP	851
QY	263	GK 264	
Db	852	GK 853	

RESULTS

```

US-10-124-905-12
;
; Sequence 12, Application US/10124905
; Patent No. US20020166136A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Anderson, Darrell R.
;
; TITLE OF INVENTION: "MONKEY MONOCLONAL ANTIBODIES SPECIFIC
;
; TITLE OF INVENTION: TO HUMAN B7.1 AND/OR B7.2 PRIMATIZED FORMS
;
; THEREOF.

```

Db 452 FSCVMHEALHNHYTKQSLSPGK 476

RESULT 8

US-09-948-429B-12

Sequence 12, Application US/09948429B

Patent No. US20020177689A1

GENERAL INFORMATION:

APPLICANT: Anderson, Darrell R.

TITLE OF INVENTION: "MONKEY MONOCLONAL ANTIBODIES SPECIFIC TO HUMAN B7.1 AND/OR B7.2 PRIMATIZED FORMS THEREOF,

TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING, AND USE THEREOF AS IMMUNOSUPPRESSANTS"

TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING, AND USE THEREOF AS IMMUNOSUPPRESSANTS"

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS

STREET: 699 Prince Street

CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22314

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/948,429B

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/383,916

FILING DATE:

APPLICATION NUMBER: US 08/487,550

FILING DATE: 07-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Teskin, Robin L.

REGISTRATION NUMBER: 35,030

REFERENCE/DOCKET NUMBER: 012712-131

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-836-6620

TELEFAX: 703-836-2021

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 476 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-948-429B-12

Query Match 88.2%; Score 1261; DB 9; Length 476;

Best Local Similarity 53.0%; Pred. No. 3.3e-32;

Matches 236; Conservative 7; Mismatches 189; Gaps 2;

Qy 9 KVAEKLK-----EAFDWLKAFYD----- 26

Db 32 KPSETLSLTCVAGSGSISGGYGWGIROPPGKGLWIGSFYSSGNTYINPLSKSQVTIS 91

Qy 27 ----- 26

Db 92 TDTSKNQFSLKNSMTAADTAVYYCYVRDLRFVSVGMVYNNWFDVWGPGLVTVSSASTKG 151

Qy 27 ----- 26

Db 152 PSVFPLAPSSKSTSGGTAALGCLVKDYFPEPTVYSWNSGALTSGVHTFPAVLQSSGLYSL 211

Qy 27 -----KVAEKLKAFMDKTHCTCPCPAPPELLGGPSVFL 59

Db 212 SSVVTPSSSLGTQTYICNVNHPKSNKTKVDDKAEPKSCDKTHCTCPAPPELLGGPSVFL 271

Qy 60 FPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYDGVGVHNAKTKPREEQNSTYRV 119

Db 272 FPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYDGVGVHNAKTKPREEQNSTYRV 331

Qy 120 VSLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQ 179

Db 332 VSLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQ 391

Qy 180 VSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFYYFYSKLTVDKSRWQQGNV 239

Db 392 VSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLDSDGSFYYFYSKLTVDKSRWQQGNV 451

Qy 240 FSCVMHEALHNHYTKQSLSPGK 264

Db 452 FSCVMHEALHNHYTKQSLSPGK 476

RESULT 9

US-09-832-659-2

Sequence 2, Application US/09832659

Patent No. US20020155547A1

GENERAL INFORMATION:

APPLICANT: BIOGEN, INC.

TITLE OF INVENTION: Interferon-Beta Fusion Proteins and Uses

FILE REFERENCE: A064PCTSEQ

CURRENT APPLICATION NUMBER: US/09/832,659

PRIOR FILING DATE: 2001-04-11

PRIOR APPLICATION NUMBER: 60/120,237

PRIOR FILING DATE: 1999-02-16

PRIOR APPLICATION NUMBER: 60/104,491

PRIOR FILING DATE: 1998-10-16

NUMBER OF SEQ ID NOS: 44

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 399

TYPE: PRT

ORGANISM: murine

US-09-832-659-2

Query Match 88.1%; Score 1260; DB 9; Length 399;

Best Local Similarity 82.9%; Pred. No. 2.3e-32;

Matches 238; Conservative 5; Mismatches 10; Indels 34; Gaps 3;

Qy 3 LKAFYDKVAEKLKE-----AFDWLKAFY-----DKVAEKLKEAFM 37

Db 122 LKRYIGRIHLHYLKAKEYSHCAWTIVRVEILRNFFINRLTGYLRNDDDDKV----- 172

Qy 38 DKHTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 97

Db 173 DKHTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 232

Qy 98 GVEVHNAKTKPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 157

Db 233 GVEVHNAKTKPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 292

Qy 158 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 217

Db 293 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 352

Qy 218 DGSFFLYSKLTVDKSRWQQGNVFCFSVMHEALHNHYTKQSLSPGK 264

Db 353 DGSFFLYSKLTVDKSRWQQGNVFCFSVMHEALHNHYTKQSLSPGK 399

RESULT 10

US-09-746-359A-54

Sequence 54, Application US/09746359A

Patent No. US20020042366A1

GENERAL INFORMATION:

APPLICANT: Thompson, Penny

APPLICANT: Foster, Donald C.

APPLICANT: Xu, Wenfeng

APPLICANT: Madden, Karen L.

APPLICANT: Kelly, James D.

APPLICANT: Sprecher, Cindy A.

APPLICANT: Blumberg, Hal

```
; APPLICANT: Eagan, Maribeth A.
; APPLICANT: Jaspers, Stephen R.
; APPLICANT: Chandrasekher, Yasmin A.
; APPLICANT: No. US20020042366Alak, Julia E.
; TITLE OF INVENTION: Method for Treating Inflammation
; FILE REFERENCE: 99-108
; CURRENT APPLICATION NUMBER: US/09/746,359A
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/171,969
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/213,341
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 54
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-746-359A-54
```

```
Query Match      88.0%; Score 1259; DB 10; Length 547;
Best Local Similarity 48.6%; Pred. No. 5.2e-32;
Matches 241; Conservative 4; Mismatches 18; Indels 233; Gaps 4;
```

```
QY 2 WLK-----AFYDKV-----10
||
Db 52 WLKSECRNINRYCDLSAETSDYEHQYAKVKAINGTKCSKWAESGRFPFLETOIGPP 111
||
QY 11 -----AEKLKE-----16
||
Db 112 EVALTTDEKSIYVLTAPKWKRNPEPLPVSQQIYSLNLYNVSLNRTKSRTWSOCVTN 171
||
QY 17 ---AFDWL-----21
||
Db 172 HTLVLTWLEPNTLYCVHVESFVPGPPRAQPSKQCAKTLKDQSSEASTKGPSVFPPLAPS 231
||
QY 22 -----21
||
Db 232 SKSTSGGTAALGCLVKDYFPEPTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVTVTPSS 291
||
QY 22 -----KAFYDKVAEKLKAFMDKTHTCPPCPAPPELLGGPSVFLFPPPKKDTL 68
||
Db 292 SLGTQYICNVNHNKPSNTKVDKKVEPKSCDKTHTCPPCPAPPELLGGPSVFLFPPPKKDTL 351
||
QY 69 MISRTPEVTCVVVDVSHEDPEVKFNKYVDGVVHNNAKTKPREEQYNSTYRVVSVLTVLHQ 128
||
Db 352 MISRTPEVTCVVVDVSHEDPEVKFNKYVDGVVHNNAKTKPREEQYNSTYRVVSVLTVLHQ 411
||
QY 129 DWLNGKEYCKKYSNKAALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKG 188
||
Db 412 DWLNGKEYCKKYSNKAALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKG 471
||
QY 189 FYPDSIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVSCSVMHEA 248
||
Db 472 FYPDSIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVSCSVMHEA 531
||
QY 249 LHNHYTQKSLSLSPGK 264
||
Db 532 LHNHYTQKSLSLSPGK 547
||
```

```
RESULT 11
US-09-746-359A-53
; Sequence 53, Application US/09746359A
; Patent No. US20020042366A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Penny
; APPLICANT: Foster, Donald C.
; APPLICANT: Xu, Wenfeng
; APPLICANT: Madden, Karen L.
; APPLICANT: Kelly, James D.
; APPLICANT: Sprecher, Cindy A.
; APPLICANT: Blumberg, Hal
```

```
; APPLICANT: Eagan, Maribeth A.
; APPLICANT: Jaspers, Stephen R.
; APPLICANT: Chandrasekher, Yasmin A.
; APPLICANT: No. US20020042366Alak, Julia E.
; TITLE OF INVENTION: Method for Treating Inflammation
; FILE REFERENCE: 99-108
; CURRENT APPLICATION NUMBER: US/09/746,359A
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/171,969
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/213,341
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 53
; LENGTH: 571
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-746-359A-53
```

```
Query Match      88.0%; Score 1259; DB 10; Length 571;
Best Local Similarity 48.6%; Pred. No. 5.8e-32;
Matches 241; Conservative 4; Mismatches 18; Indels 233; Gaps 4;
```

```
QY 2 WLK-----AFYDKV-----10
||
Db 76 WLKSECRNINRYCDLSAETSDYEHQYAKVKAINGTKCSKWAESGRFPFLETOIGPP 135
||
QY 11 -----AEKLKE-----16
||
Db 136 EVALTTDEKSIYVLTAPKWKRNPEPLPVSQQIYSLNLYNVSLNRTKSRTWSOCVTN 195
||
QY 17 ---AFDWL-----21
||
Db 196 HTLVLTWLEPNTLYCVHVESFVPGPPRAQPSKQCAKTLKDQSSEASTKGPSVFPPLAPS 255
||
QY 22 -----21
||
Db 256 SKSTSGGTAALGCLVKDYFPEPTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVTVTPSS 315
||
QY 22 -----KAFYDKVAEKLKAFMDKTHTCPPCPAPPELLGGPSVFLFPPPKKDTL 68
||
Db 316 SLGTQYICNVNHNKPSNTKVDKKVEPKSCDKTHTCPPCPAPPELLGGPSVFLFPPPKKDTL 375
||
QY 69 MISRTPEVTCVVVDVSHEDPEVKFNKYVDGVVHNNAKTKPREEQYNSTYRVVSVLTVLHQ 128
||
Db 376 MISRTPEVTCVVVDVSHEDPEVKFNKYVDGVVHNNAKTKPREEQYNSTYRVVSVLTVLHQ 435
||
QY 129 DWLNGKEYCKKYSNKAALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKG 188
||
Db 436 DWLNGKEYCKKYSNKAALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKG 495
||
QY 189 FYPDSIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVSCSVMHEA 248
||
Db 496 FYPDSIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVSCSVMHEA 555
||
QY 249 LHNHYTQKSLSLSPGK 264
||
Db 556 LHNHYTQKSLSLSPGK 571
||
```

```
RESULT 12
US-09-740-002-27
; Sequence 27, Application US/09740002
; Patent No. US20020001798A1
; GENERAL INFORMATION:
; APPLICANT: BRAMS, PETER
; APPLICANT: MORROW, PHILLIP
; TITLE OF INVENTION: NEUTRALIZING HIGH AFFINITY HUMAN MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: SPECIFIC TO RSV F-PROTEIN AND METHODS FOR THEIR
; TITLE OF INVENTION: MANUFACTURE AND THERAPEUTIC USE THEREOF
; FILE REFERENCE: 037003-0275759
; CURRENT APPLICATION NUMBER: US/09/740,002
```


; CURRENT FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 09/335,697
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 08/488,376
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-740-002-27

Query Match 88.0%; Score 1258; DB 10; Length 475;
Best Local Similarity 55.3%; Pred. No. 4e-32;
Matches 236; Conservative 9; Mismatches 12; Indels 170; Gaps 3;
QY 1 DWLKAFYDKVAEKLKEAFDWL----- 21
DB 56 NWIR-----QPPGKALEWLARIDWDDDTFYASLKTSLISKDTSKNQVVLRTNVDP 108
QY 22 -----KAFYD----- 26
DB 109 VDTATYFCARASLYDSDFLYFAYHAYWGQTVTVVSSASTKGPSVFPLAPSKSTSGGTA 168
QY 27 ----- 26
DB 169 ALGCLVKDYFPEPTVSNWNSGALTSGVHTFPAVLQSSGLYSLSSVTVVPSSSLGTQTYIC 228
QY 27 -----KVAEKLKEAFMDKTHCTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVT 77
DB 229 NVNHRKPSNTKVDKAEKPKSCDTHCTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVT 288
QY 78 CVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYK 137
DB 289 CVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYK 348
QY 138 CKVSNKALPAPIEKTISKAKGQPREPOVYITLPPSRDELTKNQVSLTCLVKGFYPSDIAVE 197
DB 349 CKVSNKALPAPIEKTISKAKGQPREPOVYITLPPSRDELTKNQVSLTCLVKGFYPSDIAVE 408
QY 198 WESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQOGNFSCSVMHAEALHNHYTQKS 257
DB 409 WESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQOGNFSCSVMHAEALHNHYTQKS 468
QY 258 LSLSPGK 264
DB 469 LSLSPGK 475

RESULT 13
US-10-091-236-17
; Sequence 17, Application US/10091236
; Patent No. US20020168360A1
; GENERAL INFORMATION:
; APPLICANT: DINGIVAN, CHRISTINE A.
; TITLE OF INVENTION: METHODS OF PREVENTING OR TREATING INFLAMMATORY OR AUTOIMMUNE
; TITLE OF INVENTION: DISORDERS BY ADMINISTERING INTEGRIN ALPHA-V-BETA-3 ANTAGONISTS
; TITLE OF INVENTION: COMBINATION WITH OTHER PROPHYLACTIC OR THERAPEUTIC AGENTS
; FILE REFERENCE: 10271-053-999
; CURRENT APPLICATION NUMBER: US/10/091,236
; CURRENT FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: US 60/273,098
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/316,321
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-091-236-17

Query Match 87.8%; Score 1256; DB 9; Length 347;
Best Local Similarity 78.5%; Pred. No. 2.3e-32;
Matches 238; Conservative 3; Mismatches 11; Indels 51; Gaps 3;
QY 2 WLKAFYDKVAEKLKEAF-----DWL 21
DB 56 WKKQ-KDKVAELEENSEFFRAFSFKNRVYLDTVYSGSLTIYNLTSSDEDEYEMESPNTIDTM 114
QY 22 KAFYDKVAEKLKEAFMDKTHCTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVV 81
DB 115 KFF-----LYVDKTHCTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVV 164
QY 82 DVSHPEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKYS 141
DB 165 DVSHPEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKYS 224
QY 142 NKALPAPIEKTISKAKGQPREPOVYITLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESN 201
DB 225 NKALPAPIEKTISKAKGQPREPOVYITLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESN 284
QY 202 GOPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQOGNFSCSVMHAEALHNHYTQKSLSLS 261
DB 285 GOPENNYKTTTPVLDSDGSFFLYSKLTVDKSRWQOGNFSCSVMHAEALHNHYTQKSLSLS 344
QY 262 PGK 264
DB 345 PGK 347

RESULT 14
US-10-091-313-7
; Sequence 7, Application US/10091313
; Publication No. US20030044406A1
; GENERAL INFORMATION:
; APPLICANT: DINGIVAN, CHRISTINE
; TITLE OF INVENTION: METHODS OF PREVENTING OR TREATING INFLAMMATORY OR AUTOIMMUNE
; TITLE OF INVENTION: DISORDERS BY ADMINISTERING CD2 ANTAGONISTS IN COMBINATION WITH
; TITLE OF INVENTION: PROPHYLACTIC OR THERAPEUTIC AGENTS
; FILE REFERENCE: 10271-063
; CURRENT APPLICATION NUMBER: US/10/091,313
; CURRENT FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: US 60/273,098
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/346,918
; PRIOR FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-091-313-7
Query Match 87.8%; Score 1256; DB 9; Length 347;
Best Local Similarity 78.5%; Pred. No. 2.3e-32;
Matches 238; Conservative 3; Mismatches 11; Indels 51; Gaps 3;
QY 2 WLKAFYDKVAEKLKEAF-----DWL 21
DB 56 WKKQ-KDKVAELEENSEFFRAFSFKNRVYLDTVYSGSLTIYNLTSSDEDEYEMESPNTIDTM 114
QY 22 KAFYDKVAEKLKEAFMDKTHCTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVV 81
DB 115 KFF-----LYVDKTHCTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVV 164
QY 82 DVSHPEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKYS 141
DB 165 DVSHPEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKYS 224
QY 142 NKALPAPIEKTISKAKGQPREPOVYITLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESN 201
DB 225 NKALPAPIEKTISKAKGQPREPOVYITLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESN 284

Qy 202 GQPNNTTTPVLDSDGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSL 261
Db 285 GQPNNTTTPVLDSDGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSL 344
Qy 262 PGK 264
Db 345 PGK 347

RESULT 15
US-10-091-268-7
; Sequence 7, Application US/10091268
; Publication No. US20030068320A1
; GENERAL INFORMATION:
; APPLICANT: DINGIVAN, CHRISTINE A
; TITLE OF INVENTION: METHODS OF ADMINISTERING/DOSING CD2 ANTAGONISTS FOR THE PREVENTION
; OF INFLAMMATORY DISORDERS
; FILE REFERENCE: 10271-054-999
; CURRENT APPLICATION NUMBER: US/10/091,268
; CURRENT FILING DATE: 2002-03-04
; PRIOR APPLICATION NUMBER: US 60/273,098
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: US 60/346,918
; PRIOR FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-091-268-7

Query Match 87.88; Score 1256; DB 9; Length 347;
Best Local Similarity 78.58; Pred. No. 2.3e-32;
Matches 238; Conservative 3; Mismatches 11; Indels 51; Gaps 3;
Qy 2 WLKAFYDKVAEKLKEAF-----DWL 21
Db 56 WKQK-KDKVAELENSEFRASFSEKRNRYLDTVSGSLTIYNTLTSSDEDEYEMESPNITDTM 114
Qy 22 KAFYDKVAEKLKEAFMDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVV 81
Db 115 KFF-----LYVDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVV 164
Qy 82 DVSHEDPEVKNWYVDGVEVHNATKPREQYNSTYRVSVLTVLDHODWLNKREYCKVS 141
Db 165 DVSHEDPEVKNWYVDGVEVHNATKPREQYNSTYRVSVLTVLDHODWLNKREYCKVS 224
Qy 142 NKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESN 201
Db 225 NKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESN 284
Qy 202 GQPNNTTTPVLDSDGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSL 261
Db 285 GQPNNTTTPVLDSDGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSL 344
Qy 262 PGK 264
Db 345 PGK 347

Search completed: April 21, 2003, 10:51:19
Job time : 14.1647 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:26 ; Search time 19.4118 Seconds
(without alignments)
1307.428 Million cell updates/sec

Title: 7LINK7LINK2

Perfect score: 1430

Sequence: 1 DWLKAFYDKVAEKLKEAFDW.....MHEALHNHYTKLSLSLSPGK 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1243	86.9	330	1 GHU	Ig gamma-1 chain C
2	1242	86.9	374	2 S69339	Ig heavy chain V r
3	1233	86.2	255	4 S31866	Ig gamma-1 chain C r
4	1190	83.2	234	2 P70207	Ig gamma-1 chain C r
5	1165	81.5	377	2 A23511	Ig gamma-3 chain C
6	1163	81.3	377	2 A60764	Ig gamma-3 chain C
7	1148	80.3	326	1 G2HU	Ig gamma-2 chain C
8	1144	80.0	327	1 G4HU	Ig gamma-4 chain C
9	1143	79.9	289	1 G3HUW1	Ig gamma-3 heavy c
10	923	64.5	323	1 GHRB	Ig gamma chain C r
11	917	64.1	328	2 I47160	Ig gamma 2b chain
12	917	64.1	328	2 I47159	Ig gamma 2a chain
13	911	63.7	277	2 I47162	Ig gamma 4 chain c
14	899	62.9	329	1 G2GP	Ig gamma-2 chain C
15	889	62.2	328	2 I47158	Ig gamma 1 chain c
16	885	61.9	328	2 I47161	Ig gamma 3 chain c
17	874	61.1	470	2 S22080	Ig heavy chain pre
18	874	61.1	472	2 S31459	Ig gamma-1 chain -
19	854	59.7	308	2 C30554	Ig heavy chain C r
20	853	59.7	329	1 G3MSC	Ig gamma-3 chain C
21	850	59.4	444	2 PC4436	Ig gamma-2b chain
22	849	59.4	333	2 PS0018	Ig gamma-3 chain C
23	842	58.9	398	1 G3MSM	Ig gamma-2a chain
24	831	58.1	469	2 S37483	Ig gamma-1 chain C
25	826	57.8	326	2 PS0017	Ig gamma-1 chain C
26	823	57.6	324	1 G4MS	Ig gamma-2a chain
27	823	57.6	446	2 S20295	Ig gamma-2c chain
28	818	57.2	329	2 S00847	Ig gamma-1 chain C
29	818	57.2	393	1 G3MSM	Ig gamma-1 chain C

30	815	57.0	330	1 G2MSA	Ig gamma-2a chain
31	810	56.6	399	1 G2MSAM	Ig gamma-2a chain
32	809	56.6	335	1 G2MSAB	Ig gamma-2a chain
33	807	56.4	474	1 G2MS11	Ig gamma-2b chain
34	795	55.6	322	2 PS0019	Ig gamma-2a chain
35	785	54.9	475	2 S01321	Ig gamma-2b chain
36	784	54.8	405	1 G2MSBM	Ig gamma-2 chain C
37	774	54.1	327	2 S06611	Ig gamma heavy cha
38	707	49.4	180	2 I46732	Ig heavy chain V-I
39	605	42.3	218	2 A36040	Ig heavy chain VHI
40	601	42.0	249	2 S69340	Ig gamma-1 chain C
41	571	39.9	152	2 S14236	titin, cardiac mus
42	485	33.9	26926	1 I38344	Ig y heavy chain (
43	444	31.0	572	2 B46529	Ig epsilon chain C
44	435	30.4	548	2 S38864	Ig heavy chain C r
45	413	28.9	448	2 S03186	Ig heavy chain C r

ALIGNMENTS

RESULT 1

GHU

Ig gamma-1 chain C region - human

C:Species: Homo sapiens (man)

C:Date: 31-Jan-1981 #sequence_revision 18-Aug-1982 #text_change 16-Jul-1999

C:Accession: A93433; S33887; B90563; A90564; B91668; A91723; A02146

R:Ellison, J.W.; Berson, B.J.; Hood, L.E.

Nucleic Acids Res. 10, 4071-4079, 1982

A:Title: The nucleotide sequence of a human immunoglobulin C-gamma gene.

A:Reference number: A93433; MUID:82274238; PMID:6287432

A:Accession: A93433

A:Molecule type: DNA

A:Residues: 1-330 <ELL>

A:Cross-references: EMBL:217370

A:Note: This sequence has the G1m(17) allotypic marker, 97-Lys, and the G1m(1) marker

A:Note: Lys-330 is removed after translation

R:Harris, L.J.

Submitted to the EMBL Data Library, October 1992

A:Reference number: S33904

A:Accession: S36861

A:Molecule type: DNA

A:Residues: 2-330 <HAR>

A:Cross-references: EMBL:217370

R:Takahashi, N.; Ueda, S.; Obata, M.; Nikaido, T.; Nakai, S.; Honjo, T.

Cell 29, 671-679, 1982

A:Title: Structure of human immunoglobulin gamma genes: implications for evolution of

A:Reference number: S33887; MUID:83001943; PMID:6811139

A:Accession: S33887

A:Molecule type: DNA

A:Residues: 88-113;235-330 <TAK>

A:Cross-references: EMBL:217370

R:Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman

Biochemistry 9, 3161-3170, 1970

A:Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid seq

A:Reference number: A90563; MUID:71064024; PMID:5489771

A:Contents: myeloma protein Eu

A:Accession: B90563

A:Molecule type: protein

A:Residues: 1-96,'R',98-135 <CUN>

A:Note: This sequence has the G1m(3) marker, 97-Arg

R:Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.

Biochemistry 9, 3171-3181, 1970

A:Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid se

A:Reference number: A90564; MUID:71064025; PMID:5530842

A:Contents: Eu

A:Accession: A90564

A:Molecule type: protein

A:Residues: 136-154,'Q',156-165,'Q',167-176,'Q',178-194,'N',196-197,'D',199-238,'E',2

A:Note: This sequence has the G1m(non-1) markers, 239-Glu and 241-Met

R:Ponstingl, H.; Hilschmann, N.

Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976

A:Title: Die Primaerstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein Ni

QY 98 GVEVHNAKTPREEQYNSTYRVVSVTLVHLDWLNKCKYKCKVSNKALPAPIEKTISKAK 157
|||||
Db 89 GVEVHNAKTPREEQYNSTYRVVSVTLVHLDWLNKCKYKCKVSNKALPAPIEKTISKAK 148
|||||
QY 158 GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 217
|||||
Db 149 GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 208
|||||
QY 218 DGSFFLYSKLTVDKSRWQOGNFVSCVMHEALHNHYTQKSLSLSPGK 264
|||||
Db 209 DGSFFLYSKLTVDKSRWQOGNFVSCVMHEALHNHYTQKSLSLSPGK 255
|||||
RESULT 4
PT0207
Ig gamma chain C region - chimpanzee
C:Species: Pan troglodytes (chimpanzee)
C:Date: 23-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 16-Jul-1999
C:Accession: PT0207
R:Ehrlich, P.H.; Moustafa, Z.A.; Oestberg, L.
Mol. Immunol. 28, 319-322, 1991
A:Title: Nucleotide sequence of chimpanzee Fc and hinge regions.
A:Reference number: PT0207; MUID:91287716; PMID:2062315
A:Accession: PT0207
A:Molecule type: mRNA
A:Residues: 1-234 <EHR>
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:48-117/Domain: immunoglobulin homology <IMM>

Query Match 83.2%; Score 1190; DB 2; Length 234;
Best Local Similarity 95.2%; Pred. No. 1.8e-32;
Matches 220; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
QY 27 KVAEKLKEAFMDKTHTCPPCAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVDVYSHE 86
|| :||: |||||
Db 4 KVDKVEPKSCDTHTCPPCAPELLGGPSVFLFPPPKDTLMISRTPEVTCVVDVYSHE 63
|||||
QY 87 DPEVKFNWYVDGVEVHNAKTPREEQYNSTYRVVSVTLVHLDWLNKCKYKCKVSNKALP 146
|||||
Db 64 DPEVKFNWYVDGVEVHNAKTPREEQYNSTYRVVSVTLVHLDWLNKCKYKCKVSNKALP 123
|||||
QY 147 APIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPEN 206
|||||
Db 124 APIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPEN 183
|||||
QY 207 NYKTTPPVLDSDGSFFLYSKLTVDKSRWQOGNFVSCVMHEALHNHYTQKS 257
|||||
Db 184 NYKTTPPVLDSDGSFFLYSKLTVDKSRWQOGNFVSCVMHEALHNHYTQKS 234
|||||

RESULT 5
A23511
Ig gamma-3 chain C region (allotype G3m(b)) - human
C:Species: Homo sapiens (man)
C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 23-Jul-1999
C:Accession: A23511
R:Huck, S.; Fort, P.; Crawford, D.H.; Lefranc, M.P.; Lefranc, G.
Nucleic Acids Res. 14, 1779-1789, 1986
A:Title: Sequence of a human immunoglobulin gamma 3 heavy chain constant region gene: c
A:Reference number: A23511; MUID:86148507; PMID:3081877
A:Accession: A23511
A:Molecule type: DNA
A:Residues: 1-377 <HUC>
A:Cross-references: GB:X03604; GB:M12958; MID:g33070; PIDN:CAA27268.1; PID:g577056
C:Genetics:
A:Gene: GDB:IGHG3
A:Cross-references: GDB:119339; OMIM:147120
A:Map position: 14q32.33-14q32.33
A:Introns: 98/3; 115/3; 130/3; 145/3; 160/3; 270/3
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:20-85/Domain: immunoglobulin homology <IMM>

Query Match 81.5%; Score 1165; DB 2; Length 377;
Best Local Similarity 76.8%; Pred. No. 6.1e-31;
Matches 218; Conservative 8; Mismatches 12; Indels 46; Gaps 2;
QY 26 DKVAEKLKEAFMDKTHT----- 42
|| | | | |
Db 95 DKRVE-LKTPLGDTHTTCPPCAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVDVYSHEDEPEVKFNWYVDGVE 153
|| | | | |
QY 43 --CPPCAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVDVYSHEDEPEVKFNWYVDGVE 100
|| | | | |
Db 154 PPCPCAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVDVYSHEDEPEVKFNWYVDGVE 213
|| | | | |
QY 101 VHNAKTPREEQYNSTYRVVSVTLVHLDWLNKCKYKCKVSNKALPAPIEKTISKAKGQP 160
|||||
Db 214 VHNAKTPREEQYNSTYRVVSVTLVHLDWLNKCKYKCKVSNKALPAPIEKTISKAKGQP 273
|||||
QY 161 REPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGS 220
|||||
Db 274 REPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGS 333
|||||
QY 221 FFYLYSKLTVDKSRWQOGNFVSCVMHEALHNHYTQKSLSLSPGK 264
|||||
Db 334 FFYLYSKLTVDKSRWQOGNFVSCVMHEALHNHYTQKSLSLSPGK 377
|||||

RESULT 6
A60764
Ig gamma-3 chain C region, form LAT - human
C:Species: Homo sapiens (man)
C:Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 16-Jul-1999
C:Accession: A60764
R:Huck, S.; Lefranc, G.; Lefranc, M.P.
Immunogenetics 30, 250-257, 1989
A:Title: A human immunoglobulin IGHG3 allele (Gmb0, bl, c3, c5, u) with an IGHG4 con
A:Reference number: A60764; MUID:90007613; PMID:2571587
A:Accession: A60764
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-377 <HUC>
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:20-85/Domain: immunoglobulin homology <IMM>

Query Match 81.3%; Score 1163; DB 2; Length 377;
Best Local Similarity 76.8%; Pred. No. 7.1e-31;
Matches 218; Conservative 8; Mismatches 12; Indels 46; Gaps 2;
QY 26 DKVAEKLKEAFMDKTHT----- 42
|| | | | |
Db 95 DKRVE-LKTPLGDTHTTCPPCAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVDVYSHEDEPEVKFNWYVDGVE 153
|| | | | |
QY 43 --CPPCAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVDVYSHEDEPEVKFNWYVDGVE 100
|| | | | |
Db 154 PPCPCAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVDVYSHEDEPEVKFNWYVDGVE 213
|| | | | |
QY 101 VHNAKTPREEQYNSTYRVVSVTLVHLDWLNKCKYKCKVSNKALPAPIEKTISKAKGQP 160
|||||
Db 214 VHNAKTPREEQYNSTYRVVSVTLVHLDWLNKCKYKCKVSNKALPAPIEKTISKAKGQP 273
|||||
QY 161 REPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGS 220
|||||
Db 274 REPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGS 333
|||||
QY 221 FFYLYSKLTVDKSRWQOGNFVSCVMHEALHNHYTQKSLSLSPGK 264
|||||
Db 334 FFYLYSKLTVDKSRWQOGNFVSCVMHEALHNHYTQKSLSLSPGK 377
|||||

RESULT 7
G2HU
Ig gamma-2 chain C region - human
C:Species: Homo sapiens (man)

Qy	14	LKEAF-----DW-----	-----LKAFYDKVAK 31
Db	29	VKDYFPEPVTVSMNSGALTSCVHTFPAVLQSSGLYSLSSVTVTPSSNFGTQTTCNV DHK 88	
Qy	32	LKEAFMDKT-----HTCCPCPAPPELLGGPSVFLFPKPKD TLMISRTPEVTCVVVDVSH 85	
Db	89	PSNTKVDKTVKCCVCCPCAPP-VAGPSVFLFPKPKD TLMISRTPEVTCVVVDVSH 147	
Qy	86	EDPEVKFNWYVDGVEVHNATKPREEQNSTYRVVSVLT VLVHODWLNGKEYCKVKSNKAL 145	
Db	148	EDPEQFNWYVDGVEVHNATKPREEQNSTFRVSVLT VLVHODWLNGKEYCKVKSNKGL 207	
Qy	146	PAPIEKTISKAKGQPREPQVY TLPSPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPE 205	
Db	208	PAPTEKITSKKGQPREPQVY TLPSPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPE 267	
Qy	206	NNYKTTTPVLDSGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 264	
Db	268	NNYKTTTPMLDSGSPFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 326	
RESULT 8			
G4HU			
Ig gamma-4 chain C region - human			
C:Species: Homo sapiens (man)			
C:Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 16-Jul-1999			
C:Accession: A90933; A90249; A02150			
R:Ellison, J.; Buxbaum, J.; Hood, L.			
DNA 1, 11-18, 1981			
A:Title: Nucleotide sequence of a human immunoglobulin C-gamma4 gene.			
A:Reference number: A90933; MUID:83157104; PMID:6299662			
A:Accession: A90933			
A:Molecule type: DNA			
A:Residues: 1-327 <ELL>			
A:Note: the sequence was determined from the germline gene			
Biochem. J. 117, 33-47, 1970			
A:Title: Human immunoglobulin subclasses. Partial amino acid sequence of			
A:Reference number: A90249; MUID:70207560; PMID:4192699			
A:Accession: A90249			
A:Molecule type: protein			
A:Residues: 1-30;81-326 <PIN>			
C:Genetics:			
A:Gene: GDB:IGHG4			
A:Cross-references: GDB:119340; OMIM:147130			
A:Map position: 14q32.33-14q32.33			
A:Introns: 99/1; 111/1; 221/1			
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical disulfide bonds. In some cases, such as IgA and IgM, the subunits associate as a dimer. In other cases, such as IgG, the subunits associate as a monomer.			
C:Superfamily: immunoglobulin C region; immunoglobulin homology			
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin			
F:20-85/Domain: immunoglobulin homology <IM1>			
F:99-110/Region: hinge			
F:134-203/Domain: immunoglobulin homology <IM2>			
F:240-307/Domain: immunoglobulin homology <IM3>			
F:14/disulfide bonds: interchain (to light chain) #status experimental			
F:27-83,141-201,247-305/disulfide bonds: #status predicted			
F:106,109/disulfide bonds: interchain (to heavy chain) #status experimental			
F:177/Binding site: carbohydrate (Asn) (covalent) #status predicted			
Query Match 80.0%; Score 1144; DB 1; Length 327;			
Best Local Similarity 84.7%; Pred. No. 1.9e-30;			
Matches 211; Conservative 9; Mismatches 6; Indels 23; Gaps			
Qy	39	KTHF-----	-----CPPCAPPELLGGPSVFLFPKPKD TLMISRTPE 75
Db	79	KTYTCNVNDHKPSNTKVDKRVESKYGPCPCPAPEFLGGPSVFLFPKPKD TLMISRTPE 138	
Qy	76	VTVCVVVDVSHDEPEVKFNWYVDGVEVHNATKPREEQYNSTYRVVSVLT VLVHODWLNGKE 135	
Db	139	VTVCVVVDVSHDEPEVFNWYVDGVEVHNATKPREEQYNSTYRVVSVLT VLVHODWLNGKE 198	
Qy	136	YKCKVSNKALPAPIEKTISKAKGQPREPQVY TLPSPSRDELTKNQVSLTCLVKGFYPSDIA 195	

Job time : 20.4718 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:21 ; Search time 10.0941 Seconds
(without alignments)
1084.766 Million cell updates/sec

Title: 7LINK7LINK2

Perfect score: 1430

Sequence: 1 DMLKAFYDKVAEKLKEAFDW.....MHEALHNNHYTKLSLSPGK 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	ID	Description
1	1243	86.9	330	1 GCL_HUMAN	P01857 homo sapien
2	1148	80.3	290	1 GC3_HUMAN	P01860 homo sapien
3	1148	80.3	326	1 GC2_HUMAN	P01859 homo sapien
4	1144	80.0	327	1 GC4_HUMAN	P01861 homo sapien
5	923	64.5	323	1 GC_RABIT	P01870 oryctolagus
6	899	62.9	329	1 GC2_CAVPO	P01862 cavia porce
7	853	59.7	329	1 GC3_MOUSE	P22436 mus musculu
8	849	59.4	333	1 GCB_RAT	P20761 rattus norv
9	842	58.9	398	1 GC3M_MOUSE	P03987 mus musculu
10	826	57.8	326	1 GCL_RAT	P20759 rattus norv
11	823	57.6	324	1 GCL_MOUSE	P01868 mus musculu
12	818	57.2	329	1 GCC_RAT	P20762 rattus norv
13	818	57.2	393	1 GCLM_MOUSE	P01869 mus musculu
14	815	57.0	330	1 GCAA_MOUSE	P01863 mus musculu
15	810	56.6	399	1 GCAM_MOUSE	P01865 mus musculu
16	809	56.6	335	1 GCAB_MOUSE	P01864 mus musculu
17	795	55.6	322	1 GCA_RAT	P20760 rattus norv
18	789	55.2	336	1 GCB_MOUSE	P01866 mus musculu
19	784	54.8	405	1 GCBM_MOUSE	P01867 mus musculu
20	389	27.2	421	1 EPC_MOUSE	P06336 mus musculu
21	386	27.0	455	1 MUC_MOUSE	P01872 mus musculu
22	383	26.8	454	1 MUC_HUMAN	P01871 homo sapien
23	382	26.7	428	1 EPC_HUMAN	P01854 homo sapien
24	381	26.6	391	1 MUCB_HUMAN	P04220 homo sapien
25	378	26.4	429	1 EPC_RAT	P01855 rattus norv
26	376	26.3	458	1 MUC_RABIT	P03988 oryctolagus
27	376	26.3	476	1 MUCM_MOUSE	P01873 mus musculu
28	366	25.6	479	1 MUCM_RABIT	P04221 oryctolagus
29	363	25.4	454	1 MUC_MESAU	P06337 mesocricetu
30	356	24.9	457	1 MUC_SUNMU	P20768 suncus muri
31	354	24.8	450	1 MUC_CANFA	P01874 canis fami
32	343	24.0	438	1 HVC2_HETER	P23085 heterodontu
33	335	23.4	438	1 HVC3_HETER	P23087 heterodontu

RESULT 1

ID	GCL_HUMAN	STANDARD;	PRT;	330 AA.
AC	P01857;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Ig gamma-1 chain C region.			
GN	IGHG1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=82274238; PubMed=6287432;			
RA	Ellison J.W., Berson B.J., Hood L.E.;			
RT	"The nucleotide sequence of a human immunoglobulin C gamma1 gene."			
RL	Nucleic Acids Res. 10:4071-4079(1982).			
RN	[2]			
RP	SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).			
RX	MEDLINE=71064024; PubMed=5489771;			
RA	Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,			
RA	Waxdal M.J., Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. VII. Amino			
RT	acid sequence of heavy-chain cyanogen bromide fragments H1-H4."			
RL	Biochemistry 9:3161-3170(1970).			
RN	[3]			
RP	SEQUENCE OF 136-329 (EU).			
RX	MEDLINE=71064025; PubMed=5530842;			
RA	Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,			
RA	Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. 8. Amino			
RT	acid sequence of heavy-chain cyanogen bromide fragments H5-H7."			
RL	Biochemistry 9:3171-3181(1970).			
RN	[4]			
RP	SEQUENCE (MYELOMA PROTEIN NIE).			
RX	MEDLINE=77070269; PubMed=826475;			
RA	Ponstingl H., Hilschmann N.;			
RT	"The rule of antibody structure. The primary structure of a			
RT	monoclonal IgG1 immunoglobulin (myeloma protein NIE). III. The			
RT	chymotryptic peptides of the H-chain, alignment of the tryptic			
RT	peptides and discussion of the complete structure."			
RL	Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).			
RN	[5]			
RP	SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.			
RX	MEDLINE=83289131; PubMed=6884994;			
RA	Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;			
RT	"Three-dimensional structure determination of antibodies. Primary			
RT	structure of crystallized monoclonal immunoglobulin IgG1 KOL, I."			
RL	Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).			
RN	[6]			
RP	DISULFIDE BONDS.			
RX	MEDLINE=71064027; PubMed=4923144;			
RA	Gall W.E., Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. X.			

34	333	23.3	6669	1	NEBU_HUMAN	P20929 homo sapien
35	327	22.9	461	1	HVCM_HETER	P23088 heterodontu
36	326	22.8	446	1	MUC_CHICK	P01875 gallus gall
37	326	22.8	6359	1	BACC_BACLI	O68008 b bacitraci
38	325	22.7	6486	1	TYCC_BACBR	O30409 b tyrocidin
39	324	22.7	393	1	HVC3_HETER	P23086 heterodontu
40	323	22.6	4967	1	RYR2_HUMAN	O92736 homo sapien
41	322	22.5	4969	1	RYR2_RABIT	P30957 oryctolagus
42	319	22.3	4568	1	DYHB_CHLRE	Q39565 chlamydomon
43	318	22.2	370	1	HVC1_HETER	P23084 heterodontu
44	318	22.2	4466	1	DYHC_TRIGR	P23098 tripneustes
45	317	22.2	4393	1	PGBW_HUMAN	P98160 homo sapien

ALIGNMENTS

RT Intrachain disulfide bonds.";
 RL Biochemistry 9:3188-3196(1970).
 RN [7]
 RP DISULFIDE BONDS.
 RX MEDLINE-77070267; PubMed-1002129;
 RA Dreker L., Schwarz J., Reichel W., Hilschmann N.;
 RT "Rule of antibody structure. The primary structure of a monoclonal
 RT IgG1 immunoglobulin (myeloma protein Nie), I: Purification and
 RT characterization of the protein, the L- and H-chains, the
 RT cyanogen bromide cleavage products, and the disulfide bridges.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
 RX MEDLINE-81208100; PubMed-7236608;
 RA Deisenhofer J.;
 RT "Crystallographic refinement and atomic models of a human Fc fragment
 RT aureus at 2.9- and 2.8-A resolution.";
 RL Biochemistry 20:2361-2370(1981).
 CC -1- MISCELLANEOUS: NIE HAS THE GIM(17) ALLOTYPIC MARKER, 97-K, & THE
 CC GIM(1) MARKERS, 239-D & 241-L. KOL & EU SEQUENCES HAVE THE GIM(3)
 CC MARKER & THE GIM (NON-1) MARKERS.
 CC -1- MISCELLANEOUS: NIE ALSO DIFFERS IN THE AMIDATION STATES OF
 CC 35,116,198,269 & 272.
 CC -1- MISCELLANEOUS: EU ALSO DIFFERS IN THE AMIDATION STATES OF RESIDUES
 CC 155, 166, 177, 195, 198, 269, AND 272 AND IN THE ORDER OF RESIDUES
 CC 268-272.
 CC -1- MISCELLANEOUS: KOL ALSO DIFFERS IN THE AMIDATION STATES OF
 CC RESIDUES 198,267&272.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; J00228; AAC82527.1; ALT_INIT.
 DR PIR; A02146; GHU.
 DR PDB; 1FC1; 15-JUL-92.
 DR PDB; 1FC2; 15-JUL-92.
 DR Genew; HGNC:5525; IGHG1.
 DR MIM; 147100; -
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003597; Ig_c1.
 DR InterPro; IPR003600; Ig_like.
 DR Pfam; PF00047; Ig; 3.
 DR SMART; SM00410; Ig_like; 1.
 DR SMART; SM00407; IGHG1; 2.
 DR PROSITE; PS00290; Ig_MHC; 2.
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
 KW 3D-structure.
 FT NON_TER 1 1
 FT DOMAIN 1 98
 FT CH1.
 FT HINGE.
 FT 99 110
 FT DOMAIN 111 223
 FT CH2.
 FT 224 330
 FT DOMAIN 224 330
 FT CH3.
 FT 27 83
 FT DISULFID 27 83
 FT 103 103
 FT DISULFID 103 103
 FT INTERCHAIN (WITH LIGHT CHAIN).
 FT 109 109
 FT DISULFID 109 109
 FT INTERCHAIN (WITH HEAVY CHAIN).
 FT 112 112
 FT DISULFID 112 112
 FT INTERCHAIN (WITH HEAVY CHAIN).
 FT 144 204
 FT DISULFID 144 204
 FT 250 308
 FT DISULFID 250 308
 FT CARBOHYD 180 180
 FT 330 330
 FT MOD_RES 330 330
 FT VARIANT 97 97
 FT 239 239
 FT VARIANT 239 239
 FT 241 241
 FT VARIANT 241 241
 FT 123 126
 FT STRAND 123 126

FT HELIX 130 134
 FT TURN 136 137
 FT STRAND 141 148
 FT STRAND 158 162
 FT TURN 163 164
 FT STRAND 165 166
 FT STRAND 175 178
 FT STRAND 183 190
 FT STRAND 193 197
 FT HELIX 198 199
 FT TURN 202 206
 FT STRAND 215 219
 FT STRAND 227 227
 FT STRAND 230 234
 FT STRAND 238 240
 FT HELIX 241 242
 FT TURN 245 256
 FT STRAND 260 266
 FT TURN 267 268
 FT STRAND 270 270
 FT STRAND 274 276
 FT STRAND 280 281
 FT TURN 283 284
 FT STRAND 287 296
 FT HELIX 297 301
 FT TURN 302 303
 FT STRAND 306 312
 FT TURN 313 314
 FT TURN 316 317
 FT STRAND 320 324
 SQ SEQUENCE 330 AA; 36106 MW; 3770EE106C2FA33D CRC64;
 Query Match 86.98; Score 1243; DB 1; Length 330;
 Best Local Similarity 96.68; Pred. No. 1.3e-35;
 Matches 230; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 Qy 27 KVAEKLEAFNDKTHTCPPCPAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVVDVSH 86
 Db 93 KVDKKVEPKSCDKTHTCPPCPAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVVDVSH 152
 Qy 87 DPEVKFNWYDGVVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALP 146
 Db 153 DPEVKFNWYDGVVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALP 212
 Qy 147 APIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPEN 206
 Db 213 APIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPEN 272
 Qy 207 NYKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNVSCFVSVMHEALHNHYTQKSLSLSPGK 264
 Db 273 NYKTTTPVLDSDGSFFLYSKLTVDKSRWQQGNVSCFVSVMHEALHNHYTQKSLSLSPGK 330
 RESULT 2
 GC3_HUMAN
 ID GC3_HUMAN STANDARD; PRT; 290 AA.
 AC P01860;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ig gamma-3 chain C region (Heavy chain disease protein) (HDC).
 GN IGHG3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE (DISEASE PROTEIN WIS).
 RX MEDLINE-81021548; PubMed-6774747;
 RA Frangione B., Rosenwasser E., Prelli F., Franklin E.C.;
 RT "Primary structure of human gamma 3 immunoglobulin deletion mutant:
 RT gamma 3 heavy-chain disease protein wis.";
 RL Biochemistry 19:4304-4308(1980).

[2]
 RN REVISIONS TO 12-97 (PROTEIN WIS).
 RX MEDLINE=71118561; PubMed=402363;
 RA Michaelson T.E., Frangione B., Franklin E.C.;
 RT "Primary structure of the 'hinge' region of human IgG3. Probable
 quadruplication of a 15-amino acid residue basic unit.";
 RL J. Biol. Chem. 252:883-889(1977).
 [3]
 RN REVISIONS TO 59-289 (PROTEIN WIS/DISEASE PROTEIN ZUC).
 RX MEDLINE=77021516; PubMed=823945;
 RA Wolfenstein-Todel C., Frangione B., Prelli F., Franklin E.C.;
 RT "The amino acid sequence of 'heavy chain disease' protein ZUC.
 Structure of the Fc fragment of immunoglobulin G3.";
 RL Biochem. Biophys. Res. Commun. 71:907-914(1976).
 [4]
 RN SEQUENCE FROM N.A. (DISEASE PROTEIN OMM).
 RX MEDLINE=82247835; PubMed=6808505;
 RA Alexander A., Steinmetz M., Barritault D., Frangione B.,
 RA Franklin E.C., Hood L., Buxbaum J.N.;
 RT "Gamma Heavy chain disease in man: cDNA sequence supports partial
 gene deletion model.";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:3260-3264(1982).
 CC -1- SUBUNIT: DIMER LINKED BY 12 DISULFIDE BONDS; IT HAS AN EXTRA
 INTERCHAIN DISULFIDE BOND AT POSITION 7 IN ADDITION TO THE 11
 NORMALLY PRESENT IN THE HINGE REGION.
 CC -1- MISCELLANEOUS: THE HEAVY CHAIN DISEASE PROTEIN WIS IS SHOWN.
 CC -1- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 42-76 WAS TAKEN FROM THE
 REF.2.
 CC -1- MISCELLANEOUS: DISEASE PROTEIN WIS IS LACKING MOST OF THE V REGION
 AND ALL OF THE CHI REGION.
 CC -1- MISCELLANEOUS: DISEASE PROTEIN ZUC LACK MOST OF THE V REGION, ALL
 OF THE CHI REGION, AND PART OF THE HINGE COMPARED WITH NORMAL
 GAMMA-3 HEAVY CHAINS.
 CC -1- MISCELLANEOUS: DISEASE PROTEIN OMM MAY REPRESENT AN ALLELIC FORM
 OR ANOTHER GAMMA CHAIN SUBCLASS.
 CC -1- MISCELLANEOUS: THE HINGE REGION IN GAMMA-3 CHAINS IS ABOUT FOUR
 TIMES AS LONG AS IN OTHER GAMMA CHAINS AND CONTAINS THREE
 IDENTICAL 15-RESIDUE SEGMENTS PRECEDED BY A SIMILAR 17-RESIDUE
 SEGMENT (12-28).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 EMBL; J00231; AAAS2805.1; ALT_SEQ.
 PIR; A02149; G3HUWI.
 HSSP; P01857; 1FC1.
 Genew; HGNC:5527; IGHG3.
 MTM; 147120; -
 InterPro; IPR003006; Ig_MHC.
 InterPro; IPR003597; Ig-cl.
 InterPro; IPR003600; Ig-like.
 Pfam; PF00047; Ig; 2.
 SMART; SM00410; Ig-like; 1.
 SMART; SM00407; IGcl; 1.
 PROSITE; PS00290; IG_MHC; 1.
 Immunoglobulin domain; Immunoglobulin C region; Glycoprotein; Repeat.
 DOMAIN 12 73
 HINGE.
 FT DOMAIN 74 183
 FT DOMAIN 184 289
 FT CH2.
 FT CH3.
 FT REPEAT 29 43
 FT REPEAT 44 58
 FT REPEAT 59 73
 FT MOD_RES 1 1
 FT CARBOHYD 6 6
 FT INTERCHAIN (WITH HEAVY CHAIN DIMER).
 FT DISULFID 7 7
 FT INTERCHAIN (WITH HEAVY CHAIN DIMER).
 FT DISULFID 24 24
 FT INTERCHAIN (WITH HEAVY CHAIN DIMER).
 FT DISULFID 27 27
 FT INTERCHAIN (WITH HEAVY CHAIN DIMER).
 FT DISULFID 33 33
 FT INTERCHAIN (WITH HEAVY CHAIN DIMER).

FT DISULFID 39 39 INTERCHAIN (WITH HEAVY CHAIN DIMER).
 FT DISULFID 42 42 INTERCHAIN (WITH HEAVY CHAIN DIMER).
 FT DISULFID 48 48 INTERCHAIN (WITH HEAVY CHAIN DIMER).
 FT DISULFID 54 54 INTERCHAIN (WITH HEAVY CHAIN DIMER).
 FT DISULFID 57 57 INTERCHAIN (WITH HEAVY CHAIN DIMER).
 FT DISULFID 63 63 INTERCHAIN (WITH HEAVY CHAIN DIMER).
 FT DISULFID 69 69 INTERCHAIN (WITH HEAVY CHAIN DIMER).
 FT DISULFID 72 72 INTERCHAIN (WITH HEAVY CHAIN DIMER).
 FT MOD_RES 140 140 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 290 290 REMOVED POST-TRANSLATIONALLY.
 FT VARIANT 126 127 OV -> EB (IN ZUC).
 FT VARIANT 134 134 /FTID=VAR_003890.
 FT VARIANT 139 139 P -> L (IN OMM).
 FT VARIANT 182 182 F -> Y (IN OMM).
 FT VARIANT 182 182 /FTID=VAR_003892.
 FT VARIANT 227 227 T -> A (IN OMM).
 FT VARIANT 227 227 /FTID=VAR_003893.
 FT VARIANT 227 227 S -> N (IN OMM).
 FT VARIANT 227 227 /FTID=VAR_003894.
 FT VARIANT 279 279 MISSING (IN ZUC).
 FT VARIANT 279 279 /FTID=VAR_003895.
 FT VARIANT 279 279 F -> Y (IN OMM).
 FT VARIANT 279 279 /FTID=VAR_003896.
 SQ SEQUENCE 290 AA; 32331 MW; E69CBC95705B2F46 CRC64;
 Query Match 80.3%; Score 1148; DB 1; Length 290;
 Best Local Similarity 74.8%; Pred. No. 1.5e-31;
 Matches 211; Conservative 15; Mismatches 11; Indels 45; Gaps 1;
 QY 28 VAEKLEAFMDKTHT----- 42
 : : : : :
 Db 9 VSSELKTLPLGDTTHTCPCEPKSCDTPPCPCPEKSCDTPPCPCPEKSCDTPPP 68
 QY 43 CPCCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 102
 : : : : :
 Db 69 CPCCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVH 128
 QY 103 NAKTKPEEYQNSTYRVVSVLTVLHQDLNGKEYCKVSNKALPAPIETKISKAKGQPRE 162
 : : : : :
 Db 129 NAKTKPEEYQNSTYRVVSVLTVLHQDLNGKEYCKVSNKALPAPIETKISKAKGQPRE 188
 QY 163 PQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSF 222
 : : : : :
 Db 189 PQVYTLPPSRDEMTKQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSF 248
 QY 223 LYSKLTVDKSRWQQGNFVSCSVMHREALNHYTQKLSLSLSPCK 264
 : : : : :
 Db 249 LYSKLTVDKSRWQQGNFVSCSVMHREALNHYTQKLSLSLSPCK 290
 RESULT 3
 GC2_HUMAN STANDARD; PRT; 326 AA.
 ID GC2_HUMAN
 AC P01859;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ig gamma-2 chain C region.
 GN IGHG2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE OF 2-326 FROM N.A.
 RX MEDLINE=82197621; PubMed=6804948;
 RA Ellison J.W., Hood L.E.;
 RT "Linkage and sequence homology of two human immunoglobulin gamma
 heavy chain constant region genes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:1984-1988(1982).
 RN [2]
 RP SEQUENCE OF 88-115 FROM N.A.

RC TISSUE-Fetal liver;
RX MEDLINE=83001943; PubMed=6811139;
RA Takahashi N., Ueda S., Obata M., Nikaigo T., Nakai S., Honjo, T.;
RT "Structure of human immunoglobulin gamma genes: implications for
RT evolution of a gene family.";
RL Cell 29:671-679(1982).
RN [3]
RP SEQUENCE OF 99-177 AND 310-326 FROM N.A.
RC TISSUE-Fetal liver;
RX MEDLINE=84235992; PubMed=6329676;
RA Krawinkel U., Rabbits T.H.;
RT "Comparison of the hinge-coding segments in human immunoglobulin gamma
RT heavy chain genes and the linkage of the gamma 2 and gamma 4 subclass
RT genes.";
RL EMBO J. 1:403-407(1982).
RN [4]
RP SEQUENCE OF 1-325 (MYELOMA PROTEIN TIL).
RX MEDLINE=81007873; PubMed=6774012;
RA Wang A.-C., Tung E., Fudenberg H.H.;
RT "The primary structure of a human IgG2 heavy chain: genetic,
RT evolutionary, and functional implications.";
RL J. Immunol. 125:1048-1054(1980).
RN [5]
RP SEQUENCE OF 1-85 AND 132-325 (MYELOMA PROTEIN ZIE).
RX MEDLINE=80001357; PubMed=113060;
RA Connell G.E., Parr D.M., Hofmann T.;
RT "The amino acid sequences of the three heavy chain constant region
RT domains of a human IgG2 myeloma protein.";
RL Can. J. Biochem. 57:758-767(1979).
RN [6]
RP SEQUENCE OF 238-275 (ZIE).
RX MEDLINE=80114419; PubMed=118920;
RA Hofmann T., Parr D.M.;
RT "A note of the amino acid sequence of residues 381-391 of human
RT immunoglobulin gamma chains.";
RL Mol. Immunol. 16:923-925(1979).
RN [7]
RP REVISIONS TO 25; 59; 60 AND 264-268 (ZIE).
RA Hofmann T., Parr D.M.;
RT Submitted (MAR-1980) to the PIR data bank.
RN [8]
RP SEQUENCE OF 1-121 (DOT).
RX MEDLINE=95255298; PubMed=7737190;
RA Stoppini M., Bellotti V., Negri A., Merlini G., Garver F., Ferri G.;
RT "Characterization of the two unique human anti-flavon monoclonal
RT immunoglobulins.";
RL Eur. J. Biochem. 228:886-893(1995).
RN [9]
RP DISULFIDE BONDS.
RX MEDLINE=72033500; PubMed=4940472;
RA Milstein C., Frangione B.;
RT "Disulphide bridges of the heavy chain of human immunoglobulin G2.";
RL Biochem. J. 121:217-225(1971).
RN [10]
RP DISULFIDE BONDS.
RX MEDLINE=69064124; PubMed=5782707;
RA Frangione B., Milstein C., Pink J.R.L.;
RT "Structural studies of immunoglobulin G.";
RL Nature 221:145-148(1969).
RN [11]
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb.ch).
RN [12]
DR EMBL: J00230; AAB59393.1; -
DR RA: A02148; G2HU.
DR HSSP: P01857; 1FC1.
DR Genew: HGNC:5526; IGHG2.
DR MIM: 147110; -

DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; Ig_like; 1.
DR SMART; SM00407; IGC1; 2.
DR PROSITE; PS00290; IG_MHC; 2.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DOMAIN 1 98
FT DOMAIN 99 110
FT DOMAIN 111 219
FT DOMAIN 220 326
FT DISULFID 14 14
FT DISULFID 27 83
FT DISULFID 102 102
FT DISULFID 103 103
FT DISULFID 106 106
FT DISULFID 109 109
FT DISULFID 140 200
FT DISULFID 246 304
FT SITE 156 156
FT MOD_RES 326 326
FT VARIANT 60 60
FT CONFLICT 109 109
FT SEQUENCE 326 AA; 35884 MW; 8310878C6878CF9C CRC64;
SQ
Query Match 80.3%; Score 1148; DB 1; Length 326;
Best Local Similarity 72.6%; Pred. No. 2.le-31;
Matches 217; Conservative 13; Mismatches 20; Indels 49; Gaps 4;
QY 14 LKEAF-----DW-----LKAFFDKVAEK 31
DB 29 VKDYPEPVTVSWNSGALTSVHTFPVAVLQSSGLYSLSSVTVVPSNFTQYTCNVDHK 88
QY 32 LKEAFMDKT-----HTCPPCPAPPELLGGPSVFLFPPPKPKDTLMISRTPEVTCVVVDVSH 85
DB 89 PSNTKVDKTVVERKCCVCEPCPAPP-VAGPSVFLFPPPKPKDTLMISRTPEVTCVVVDVSH 147
QY 86 EDPEVKFNWYDGVGVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKCKVSNKAL 145
DB 148 EDPEVQFNWYDGVGVHNAKTPREEQFNSTYRVVSVLTVLHHQDLNGKEYCKCKVSNKGL 207
QY 146 PAPIETKTISKAGQPREPOVYTLPPSRDLTKNQVSLTCLVKGFYPSDIAVWESNGQPE 205
DB 208 PAPIETKTISKAGQPREFOVYTLPPSRDEMTKNQVSLTCLVKGFYPSDIAVWESNGQPE 267
QY 206 NNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVSCSVMHREALHNHYTQKSLSLSPGK 264
DB 268 NNYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVSCSVMHREALHNHYTQKSLSLSPGK 326
RESULT 4
GC4_HUMAN STANDARD; PRT; 327 AA.
AC P01861;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
GN IGHG4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83157104; PubMed=6299662;
RA Ellison J.W., Buxbaum J.N., Hood L.E.;
RT "Nucleotide sequence of a human immunoglobulin C gamma 4 gene.";
RL DNA 1:11-18(1981).
RN [2]

```

RP SEQUENCE OF 1-30 AND 81-326.
RX MEDLINE=70207560; PubMed=4192699;
RA Pink J.R.L., Buttery S.H., de Vries G.M., Milstein C.;
RT "Human immunoglobulin subclasses. Partial amino acid sequence of the
RL constant region of a gamma 4 chain.";
RL Biochem. J. 117:33-47(1970).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: K01316; AAB59394.1; ALT_INIT.
DR PIR: A02150; G4HU.
DR HSP: P01842; 7FAB.
DR Genew; HGNC:5528; IGHG4.
DR MIM; 147130;
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_cl.
DR InterPro: IPR003600; Ig_like.
DR Pfam: PF00047; Ig; 3.
DR SMART; SM00410; Ig_like; 1.
DR SMART; SM00407; IGcl; 2.
DR PROSITE; PS00290; IG_MHC; 2.
DR Immunoglobulin domain; Immunoglobulin C region.
KW NON_TER 1
FT DOMAIN 1 98 CH1.
FT DOMAIN 99 110 HINGE.
FT DOMAIN 111 220 CH2.
FT DOMAIN 221 327 CH3.
FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 83
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 141 201
FT DISULFID 247 305
SQ SEQUENCE 327 AA; 35940 MW; 3EDBD811EF208E7A CRC64;

Query Match 80.0%; Score 1144; DB 1; Length 327;
Best Local Similarity 84.7%; Pred. No. 2.9e-31;
Matches 211; Conservative 9; Mismatches 6; Indels 23; Gaps 1;

QY 39 KTHP-----CPCPAPELLGGPSVFLFPPKPKDTLMISRTPE 75
Db 79 KTYCNVDHPSNTKVDKRVESKYGPPCPAPEFLGGPSVFLFPPKPKDTLMISRTPE 138
QY 76 VTCVVDVSHEDPEVKFNWYDGVGVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKE 135
Db 139 VTCVVDVDSQEDPEVQFNWYDGVGVHNAKTKPREQFNSTYRVVSVLTVLHQDWLNGKE 198
QY 136 YKCVSNKALPAPLEKIISAKAGPREPOVYTLPPSRDELTKNOVSLTCLVKGYPSDIA 195
Db 199 YKCVSNKGLPSLSEKTIISAKAGPREPOVYTLPPSQEEMTKNOVSLTCLVKGYPSDIA 258
QY 196 VEWESNGOPENNYKTTTPVLDSGSPFLYSLKLTVDKSRWQGNVFSQSVMEALHNHYTQ 255
Db 259 VEWESNGOPENNYKTTTPVLDSGSPFLYSLKLTVDKSRWQGNVFSQSVMEALHNHYTQ 318
QY 256 KSLSLSPCK 264
Db 319 KSLSLSLGK 327

RESULT 5
GC_RABIT STANDARD; PRT; 323 AA.
AC P01870;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)

```

```

DE Ig gamma chain C region.
OC Oryctolagus cuniculus (Rabbit).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84030930; PubMed=6313520;
RA Bernstein K.E., Alexander C.B., Mage R.G.;
RT "Nucleotide sequence of a rabbit IgG heavy chain from the recombinant
RL F-1 haplotype.";
RL Immunogenetics 18:387-397(1983).
RN [2]
RP SEQUENCE OF 1-128.
RX MEDLINE=76135469; PubMed=1243651;
RA Pratt D.M., Moile L.E.;
RT "Sequence studies on the constant region of the Fd sections of rabbit
RL immunoglobulin G of different allotype.";
RL Biochem. J. 151:337-349(1975).
RN [3]
RP SEQUENCE OF 88-266 FROM N.A.
RX MEDLINE=83299917; PubMed=6193512;
RA Martens C.L., Moore K.W., Steinmetz M., Hood L., Knight K.L.;
RT "Heavy chain genes of rabbit IgG: isolation of a cDNA encoding gamma
RL heavy chain and identification of two genomic C gamma genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:6018-6022(1982).
RN [4]
RP SEQUENCE OF 132-161.
RX MEDLINE=70110015; PubMed=5461106;
RA Fruchter R.G., Jackson S.A., Moile L.E., Porter R.R.;
RT "Sequence studies of the Fd section of the heavy chain of rabbit
RL immunoglobulin G.";
RL Biochem. J. 116:249-259(1970).
RN [5]
RP SEQUENCE OF 129-131 AND 155-322.
RA Hill R.L., Lebovitz H.E., Fellows R.E. Jr., Delaney R.;
RL (In) Killander J. (eds.);
RL Gamma globulins, Nobel symp. 3, pp.109-127, Almqvist and Wiksell,
Stockholm (1967).
CC -1- MISCELLANEOUS: REF.1 SEQUENCE HAS THE D12 ALLOTYPIC MARKER,
CC 104-THR, AND THE E14 MARKER, 185-THR. REF.3 HAS THE D11 AND E15
CC MARKERS AND REF.5 THE E15 MARKER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M16426; AAA31289.1;
DR PIR: A02161; GHRB.
DR HSP: P01857; IFCl.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_cl.
DR Pfam: PF00047; Ig; 2.
DR SMART; SM00407; IGcl; 2.
DR PROSITE; PS00290; IG_MHC; 1.
DR Immunoglobulin domain; Immunoglobulin C region.
KW NON_TER 1
FT VARIANT 104 104 T -> M (IN D11 MARKER).
FT VARIANT 185 185 T -> A (IN E15 MARKER).
FT CONFLICT 48 48 N -> E (IN REF. 2).
FT CONFLICT 71 71 V -> VPV (IN REF. 2).
FT CONFLICT 144 144 Q -> E (IN REF. 3 AND 4).
FT CONFLICT 173 173 N -> D (IN REF. 5).
FT CONFLICT 187 187 Q -> E (IN REF. 3 AND 5).
FT CONFLICT 201 201 N -> D (IN REF. 5).
FT CONFLICT 218 218 Q -> E (IN REF. 5).
FT CONFLICT 233 233 E -> Q (IN REF. 5).
FT CONFLICT 246 246 N -> D (IN REF. 5).
FT CONFLICT 256 256 E -> G (IN REF. 5).

```

```

FT CONFLICT 260 260 N -> D (IN REF. 5).
FT CONFLICT 266 266 N -> D (IN REF. 5).
FT CONFLICT 280 280 Y -> W (IN REF. 5).
FT CONFLICT 284 284 N -> S (IN REF. 5).
SQ SEQUENCE 323 AA: 35404 MW: 59E8AA118D579A8B CRC64;

Query Match 64.5%; Score 923; DB 1; Length 323;
Best Local Similarity 69.8%; Pred. No. 8.4e-24;
Matches 169; Conservative 29; Mismatches 39; Indels 5; Gaps 2;

QY 28 VAEKLKEAFMDKT---HTC--PPCAPPELLGGPSVLFPPPKPKDTLMISRTPEVTCVVD 82
Db 82 VAHPATNKVDKTVAPSTCKPTCPPELLGGPSVFIFFPKPKDTLMISRTPEVTCVVD 141

QY 83 VSHEDPEKFNWVGVGEVHNATKPREQYNSTYRVSVLTVLHODWLGKREYCKVSN 142
Db 142 VSQDDEVOFTWVINEQVETARPPREQOQFNSTIRVSTPLTHODWLGKREYCKVHN 201

QY 143 KALPAPIEKTISKAKQPREQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNG 202
Db 202 KALPAPIEKTISKARGOPLKPVYTMGPPREELSSRSVSLTCMNGFYPSDISVEWENKNG 261

QY 203 QPENNYKTPPVLDSDGSGFLSKLVDSKRWQGNVFCVSMVHEALHNYHTQKSLSP 262
Db 262 KAEDNYKTPPAVLDSGSGFLYLNKLSVPTSEWQSGDVFCTCSMVHEALHNYHTQKSLSP 321

QY 263 GK 264
Db 322 GK 323

RESULT 6
GC2_CAVPO STANDARD; PRT; 329 AA.
AC P01862;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2 chain C region.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE OF 1-3.
RA Trischmann T.M.;
RL Submitted (APR-1975) to the PIR data bank.
RN [2]
RP SEQUENCE OF 4-68.
RX MEDLINE=71058471; PubMed=5538606;
RA Birshstein B.K., Hussain O.Z., Cebra J.J.;
RT "Structure of heavy chain from strain 13 guinea pig
immunoglobulin-G(2). 3. Amino acid sequence of the region around the
half-cysteine joining heavy and light chains.";
RL Biochemistry 10:18-25(1971).
RN [3]
RP SEQUENCE OF 69-133 AND 312-329.
RX MEDLINE=71058486; PubMed=5538616;
RA Turner K.J., Cebra J.J.;
RT "Structure of heavy chain from strain 13 guinea pig
immunoglobulin-G(2). II. Amino acid sequence of the carboxyl-terminal
and hinge region cyanogen bromide fragments.";
RL Biochemistry 10:9-17(1971).
RN [4]
RP SEQUENCE OF 134-226.
RX MEDLINE=75036072; PubMed=4429665;
RA Tracey D.E., Cebra J.J.;
RT "Primary structure of the CH2 homology region from guinea pig IgG2
antibodies.";
RL Biochemistry 13:4796-4803(1974).
RN [5]
RP SEQUENCE OF 227-311.
RX MEDLINE=75036073; PubMed=4609467;

```

```

RA Trischmann T.M., Cebra J.J.;
RT "Primary structure of the CH3 homology region from guinea pig IgG2
antibodies.";
RL Biochemistry 13:4804-4811(1974).
RN [6]
RP DISULFIDE BONDS.
RX MEDLINE=71058474; PubMed=4922544;
RA Oliveira B., Lamm M.E.;
RT "Interchain disulfide bridges of guinea pig gamma-2-immunoglobulin.";
RL Biochemistry 10:26-31(1971);
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM POOLED SERUM OF STRAIN
CC 13 INBRED GUINEA PIGS.
DR PIR; A02151; G2GP.
DR HSP; P01842; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig-CL.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00410; IG_Like; 1.
DR SMART; SM00407; IGc1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1 1
FT DISULFID 16 16 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 28 79
FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 142 202
FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .).
FT DISULFID 248 308
SQ SEQUENCE 329 AA: 36074 MW: 5D231B7164D1FBA9 CRC64;

Query Match 62.9%; Score 899; DB 1; Length 329;
Best Local Similarity 67.1%; Pred. No. 5.8e-23;
Matches 167; Conservative 25; Mismatches 44; Indels 13; Gaps 2;

QY 28 VAEKLKEAFMDKT-----HTCCPPCAPPELLGGPSVLFPPPKPKDTLMISRTPEV 76
Db 81 VAHPASTTKVDKTVETPIPTZPBCTCKPPENLGGPSVFIFFPKPKDTLMISRTPRV 140

QY 77 TCVVVDVSHEDPEVKFNWYGVGEVHNATKPREQYNSTYRVSVLTVLHODWLGKREY 136
Db 141 TCVVVDVSDQDEPEVQFTWFDNKNPVGNAETKPRVEQYNTPRVESVLPQHODWLGKREY 200

QY 137 KCKVSNKALPAPIEKTISKAKQPREQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAV 196
Db 201 KCKVYNKALPAPIEKTISKAKGAPRPMDVYTLPPSRDELTKNOVSLTCLVIFPPADIAHV 260

QY 197 EWESNGQP--ENNYKTTPTPLVLDSDGSGFLYSLTVDKSRWQGNVFCVSMVHEALHNYHT 254
Db 261 EWASNRVPVSEKEYKNTPTPIEDADGSGFLYSLTVDKSNWDQGVYTCVSMVHEALHNYHT 320

QY 255 QKSLSLSPG 263
Db 321 OKAISRSPPG 329

RESULT 7
GC3_MOUSE
ID GC3_MOUSE STANDARD; PRT; 329 AA.
AC P22436;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig gamma-3 chain C region, secreted form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;

```



```

RX MEDLINE=84041483; PubMed=6314258;
RA Komaromy M., Clayton L., Rogers J., Robertson S., Kettman J.,
RA Wall R.;
RT "The structure of the mouse immunoglobulin in gamma 3 membrane gene
segment.";
RL Nucleic Acids Res. 11:6775-6785(1983).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00451; AAB59655.1;
DR EMBL; V01526; CAA24767.1; ALT_SEQ.
DR PIR; A02155; G3MSM.
DR HSP; P01857; LFC1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; Ig_like; 1.
DR SMART; SM00407; IGC1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1
FT DOMAIN 1 97
FT DOMAIN 98 113 CH1.
FT DOMAIN 114 223 HINGE.
FT DOMAIN 224 327 CH2.
FT DOMAIN 328 362 CH3.
FT TRANSMEM 346 362
FT DOMAIN 363 398
FT DOMAIN 399 432
FT CONFLICT 333 342 E -> G (IN REF. 2).
FT CONFLICT 342 342 E -> Q (IN REF. 2).
FT CONFLICT 388 388 P -> F (IN REF. 2).
SQ SEQUENCE 398 AA; 43929 MW; CF7F264B50A41B95 CRC64;

Query Match 58.9%; Score 842; DB 1; Length 398;
Best Local Similarity 52.7%; Pred. No. 8.9e-21;
Matches 134; Conservative 37; Mismatches 46; Indels 55; Gaps 4;

QY 10 VAELKKEAF-----DWLKAFYDKVAEKLKAEAFMDKTHTC----- 43
DQ 52 VSSVLQSGFSLSSLVTPSSW-----PSQIVCNVAHPASKTELIR 95
QY 44 -----PP---CPAPELLGGPSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEV 90
DQ 96 IEPRIPKPTTPGSSCPPGNILGGPSVFLPPKPKDMLISLTPKTCVVDVSEDDPDV 155
QY 91 KFNWYDGVGVHNAKTPREEQYNSTYRVSVLTVLHODWLNKKEYCKVSNKALPAPTE 150
DQ 156 HVSFWFVNKEVHTAWTPQREAOQNTFRVVSALPIQHDQMRGKFCCKVNNKALPAPTE 215
QY 151 KTISKAGQPREQVYTLPPSRDLKQVSLTCLVKGFPSDIAVWESNGQPNENYKT 210
DQ 216 RTISKPKGRAQTPQVYTIPTPPQMSKKVSLTCLVTFNSEAISVWERNGELEQDYKN 275
QY 211 TTPVLDSGDSFFLYSKLTVDKNSWQGNVFCVSVHMAHNNHYTKSLSP 262
DQ 276 TTPILSDGTIFYLYSLKTVTDTSWLGQEIFTCVSVHMAHNNHYTKSLSP 327

RESULT 10
GCL_RAT
ID GCL_RAT
AC P20759;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-1 chain C region.
```

```

OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
DR PIR; PS0017; PS0017.
DR HSP; P01842; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; IGC1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1
FT DOMAIN 1 97
FT DOMAIN 98 112 CH1.
FT DOMAIN 113 219 HINGE.
FT DOMAIN 220 326 CH2.
FT DOMAIN 327 362 CH3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 140 200
FT DISULFID 246 304
FT CARBOHYD 176 176
SQ SEQUENCE 326 AA; 35946 MW; 013BAB45EF49B9DA CRC64;

Query Match 57.8%; Score 826; DB 1; Length 326;
Best Local Similarity 43.8%; Pred. No. 1.7e-20;
Matches 145; Conservative 45; Mismatches 39; Indels 102; Gaps 3;

QY 3 LKAFY----- 7
DQ 29 VKGYPEPVTVNSGALSSGVHTPPAVLQSLGTLTSSVTPSSWTPQVTCNVAPHA 88
QY 8 --DKVAELKAEAFMDLKAFYDKVAEKLKAEAFMDKTHTCPPAPELGG----- 54
DQ 89 SSTKVDKKT-----VPRNCGGDKCPICITGSE 115
QY 55 -PSVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKNWYDGVGVHNAKTPREEQY 113
DQ 116 VSSVFLFPKPKDVLITLTPKTCVVDVSHEDPEVKNWYDGVGVHNAKTPREEQY 175
QY 114 NSTYRVSVSLTVLHODWLNKKEYCKVSNKALPAPTEKTISKAKGQPREQVYTLPPSRD 173
DQ 176 NSTERSVSELPILHODWLNKGRFRCKVTSAAPSPLEKTIKPEGRTPVHYVMSPTKE 235
QY 174 ELTKNQVSLTCLVKGFPSDIAVWESNGQPNENYKTTPPVLDSDGDSFFLYSKLTVDKSR 233
DQ 236 EMTQNEVSITCMVKGFYPPDIYVWQMGQPNENYKNTPTMTDGSYFLYSLKLVNKKK 295
QY 234 WQGNVFCVSVHMAHNNHYTKSLSPCK 264
DQ 296 WQGNVFCVSVHMAHNNHYTKSLSPCK 326

RESULT 11
GCL_MOUSE
ID GCL_MOUSE
AC P01868;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ig gamma-1 chain C region.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=80045036; PubMed=115593;
 RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
 RA Takahashi N., Mano Y.;
 RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
 RT gamma 1 chain gene."
 RL Cell 18:559-568(1979).
 RN [2]
 RP SEQUENCE OF 76-324 FROM N.A. (MYELOMA PROTEIN MOPC 31C).
 RX MEDLINE=80202559; PubMed=6769752;
 RA Obata M., Yamawaki-Kataoka Y., Takahashi N., Kataoka T., Shimizu A.,
 RA Mano Y., Seidman J.G., Peterlin B.M., Leder P., Honjo T.;
 RT "Immunoglobulin gamma 1 heavy chain gene: structural gene sequences
 RT cloned in a bacterial plasmid."
 RL Gene 9:87-97(1980).
 RN [3]
 RP SEQUENCE OF 70-322 FROM N.A. (MYELOMA PROTEIN MOPC 21).
 RX MEDLINE=80012837; PubMed=113776;
 RA Rogers J., Clarke P., Salser W.;
 RT "Sequence analysis of cloned cDNA encoding part of an immunoglobulin
 RT heavy chain."
 RL Nucleic Acids Res. 6:3305-3321(1979).
 RN [4]
 RP SEQUENCE (MYELOMA PROTEIN MOPC 21).
 RX MEDLINE=78242288; PubMed=98524;
 RA Adetugbo K.;
 RT "Evolution of immunoglobulin subclasses. Primary structure of a
 RT murine myeloma gamma chain."
 RL J. Biol. Chem. 253:6068-6075(1978).
 RN [5]
 RP DISULFIDE BONDS (MOPC 21).
 RX MEDLINE=73008889; PubMed=5073237;
 RA Svasti J., Milstein C.;
 RT "The disulphide bridges of a mouse immunoglobulin G1 protein."
 RL Biochem. J. 126:837-850(1972).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; V00793; CAA24172.1; -
 DR EMBL; V00793; CAA24173.1; -
 DR EMBL; V00793; CAA24174.1; -
 DR EMBL; V00793; CAA24175.1; -
 DR EMBL; V00795; CAA24176.1; -
 DR PIR; A02159; GIMS
 DR HSP; P01842; 7FAB.
 DR GlycosuiteDB; P01868; -
 DR MGD; MGI:96446; Igh-4.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003597; Ig_cl.
 DR Pfam; PF00047; Igh; 3.
 DR SMART; SM00407; Igh1; 2.
 DR PROSITE; P500290; IG_MHC; 1.
 DR Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
 KW Alternative splicing.
 FT NON_TER 1 1
 FT DOMAIN 1 97
 FT HINGE. 98 110
 FT DOMAIN 111 217
 FT CH2. 111 217
 FT CH3. 218 324
 FT DISULFID 27 82
 FT INTERCHAIN (WITH A LIGHT CHAIN). 102 102
 FT DISULFID 104 104
 FT INTERCHAIN (WITH A HEAVY CHAIN). 107 107
 FT DISULFID 107 107
 FT INTERCHAIN (WITH A HEAVY CHAIN). 109 109
 FT DISULFID 109 109
 FT INTERCHAIN (WITH A HEAVY CHAIN). 138 138
 FT DISULFID 138 138

FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .).
 FT DISULFID 244 302 /FTIQ-CAR_000055.
 FT MOD_RES 324 324 REMOVED POST-TRANSLATIONALLY.
 FT CONFLICT 276 276 N -> D (IN REF. 3).
 FT CONFLICT 278 278 N -> D (IN REF. 3).
 SQ SEQUENCE 324 AA; 35704 MW; A338812F3D1F2C93 CRC64;
 Query Match 57.6%; Score 823; DB 1; Length 324;
 Best Local Similarity 44.9%; Pred. No. 2e-20;
 Matches 144; Conservative 51; Mismatches 42; Indels 84; Gaps 4;
 QY 3 LKAFY-----
 Db 29 VGYFPEPTVTWNSGSLSSGVHTFPVAVLQSLDLYTLSSSVTVPSRPSETVTCNVAHDA 88
 QY 8 --DKVAEKLKEAFDWLKAFLKAEAFMDKTHTCPP--CPAPELLGGPSVFLFPPK 63
 Db 89 SSTKVDKKI-----VPRDGGCKPCICTVPEV---SSVFIFPPK 123
 QY 64 PKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVGVHNAKTKPREQVNSTYRVSVL 123
 Db 124 PRDVLITLTPKTCVVDVSHEDPEVKFNWYVDGVGVHNAKTKPREQVNSTYRVSVL 183
 QY 124 TVLHQLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLT 183
 Db 184 PIMHQLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLT 243
 QY 184 CLVKGYPYSDIAVWESNGQPNKYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVPSCS 243
 Db 244 CMITDFPEDITVWQWNGQPAENYKNTQPIINTNGSYFYSLKLVNQSWEAGNFTTCS 303
 QY 244 VMHEALHNHYTKSLSLSPGK 264
 Db 304 VLHEGLNHHHTKSLSHSPGK 324
 RESULT 12
 GCC_RAT
 ID GCC_RAT STANDARD; PRT; 329 AA.
 AC P20762;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig gamma-2C chain C region.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88166903; PubMed=3127222;
 RA Brueggemann M., Delmastro-Galfre P., Waldmann H., Calabi F.;
 RT "Sequence of a rat immunoglobulin gamma 2c heavy chain constant
 RT region cDNA: extensive homology to mouse gamma 3."
 RL Eur. J. Immunol. 18:317-319(1988).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X07189; CAA30169.1; -
 DR PIR; S00847; S00847.
 DR HSP; P01842; 7FAB.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003597; Ig-cl.
 DR InterPro; IPR003600; Ig_like.
 DR Pfam; PF00047; Igh; 2.
 DR SMART; SM00410; Ig_like; 1.

```

DR SMART; SM00407; IGC1; 2.
KW PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DOMAIN 1 97
FT DOMAIN 98 113
FT DOMAIN 114 222
FT DOMAIN 223 329
FT DISULFID 15 15
FT DISULFID 27 82
FT DISULFID 111 111
FT DISULFID 113 113
FT DISULFID 143 203
FT DISULFID 249 307
SQ SEQUENCE 329 AA; 36571 MW; 5FCD7B7933850773 CRC64;

Query Match 57.2%; Score 818; DB 1; Length 329;
Best Local Similarity 57.9%; Pred. No. 3.2e-20;
Matches 146; Conservative 42; Mismatches 38; Indels 26; Gaps 2;

QY 39 KTHTC-----PP---CPAPELLGGPSVFLPPPKPKDTLMISR 72
DB 78 QIVTCSVAHPATKSNLRIKRIEPRRPKPRPTDICSNDLGRSPVIFPPPKDILMLIL 137
QY 73 TPEVTGVVVDVSHDEPKVFNVDGVEVHNAKTKPREEQYNSTYRVSVLTVLHQDWLN 132
DB 138 TPKVTCVVVDVSEEDPVQFWSFVDNRVFTAQTPHEEQNLGTRVYSTLHIHQDWNS 197
QY 133 GKEYCKVSNKALPAPIETISKAGOPREPQVITLPSRDELTKNQVSLCLVKGFPYS 192
DB 198 GKEFKCKVNNKDLSPIEKTIKPKGKARTPOVYITPPPREQMSKNKVSLSCTWYTSFYPA 257
QY 193 DIAVEWESNGPENNKYKTPPVLDSGDSFFLYSKLTVKSRWQGNVFCSCVMHEALHNH 252
DB 258 SISVEWERNGELEQDYKNTLPVLDSDESIFYLSKLSVDTDSWNRGDIYTCVWHEALHNH 317
QY 253 YTKSLSLSPGK 264
DB 318 HTQKNLSRSPGK 329

RESULT 13
GCIM_MOUSE STANDARD; PRT; 393 AA.
AC P01869;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE I9 gamma-1 chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=80045036; PubMed=115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
RT gamma 1 chain gene.";
RL Cell 18:559-568(1979).
[2]
RP SEQUENCE OF 323-393 FROM N.A.
RX MEDLINE=82197626; PubMed=6804950;
RA Tyler B.M., Cowman A.F., Gerondakis S.D., Adams J.M., Bernard O.;
RT "mRNA for surface immunoglobulin gamma chains encodes a highly
RT conserved transmembrane sequence and a 28-residue intracellular
RT domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2008-2012(1982).
[3]
RP SEQUENCE OF 323-366 FROM N.A.
RX MEDLINE=82115295; PubMed=6799207;
RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,

```

```

RA Eisenberg D., Wall R.;
RT "Gene segments encoding transmembrane carboxyl termini of
RT immunoglobulin gamma chains.";
RL Cell 26:19-27(1981).
[4]
RP SEQUENCE OF 1-44 FROM N.A.
RX MEDLINE=82222190; PubMed=6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
RT immunoglobulin gamma chains.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
CC -1- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA
CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED-
CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
CC BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
CC IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
CC SEGMENT OF MU CHAINS.
-----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; V00793; CAA24172.1; -
DR EMBL; V00793; CAA24173.1; -
DR EMBL; V00793; CAA24174.1; -
DR PIR; B02159; GIMSM.
DR HSP; P01842; 7FAB.
DR MGI; 96446; Igh-4.
DR InterPro; IPR003006; Iq_MHC.
DR InterPro; IPR003597; Iq-cl.
DR Pfam; PF00047; Iq; 3.
DR SMART; SM00407; IGC1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing; Transmembrane.
FT NON_TER 1
FT DOMAIN 1 97
FT DOMAIN 98 110
FT DOMAIN 111 217
FT DOMAIN 218 324
FT DISULFID 27 82
FT DISULFID 102 102
FT DISULFID 104 104
FT DISULFID 107 107
FT DISULFID 109 109
FT DISULFID 138 198
FT CARBOHYD 174 174
FT DISULFID 244 302
FT TRANSMEM 340 357
FT DOMAIN 358 393
SQ SEQUENCE 393 AA; 43386 MW; 4CC88343B7A1CE27 CRC64;

Query Match 57.2%; Score 818; DB 1; Length 393;
Best Local Similarity 59.6%; Pred. No. 5.5e-20;
Matches 140; Conservative 48; Mismatches 42; Indels 5; Gaps 2;

QY 31 KLKEAFMDKTHTCPP--CPAPELLGGPSVFLPPPKPKDTLMISRPEVTCVVVDVSHEDP 88
DB 92 KVDKKIVPRDCGCKPCICTPEV---SSVFFPPKPKDVLITLTPKVTVCVVVDISKDDP 148
QY 89 EVKENWYVDGVEVINAKTKPREEQYNSTYRVSVLTVLHQDWLNGKEYCKVSNKALPAP 148
DB 149 EVQFSWFDVDDVEVHTAQTQPREQFNSTFRSVSELPIMHQDWLNGKFKCVNSAAPPAP 208
QY 149 IEKTIKAGOPREPQVITLPPSPRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNY 208
DB 209 IEKTIKTKGRPKAPQVYITPPPKQAKDKVSLTCTITDFFPDITVEWQWNGQPAENY 268
QY 209 KTTPPVLDSGDSFFLYSKLTVKSRWQGNVFCSCVMHEALHNHHTYKSLSPG 263

```


Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	1256	87.8	471	4	Q8TC77	Q8tc77	homo sapien
2	1251	87.5	701	4	Q96P08	Q96pq8	homo sapien
3	1155	80.8	473	4	Q8TC63	Q8tc63	homo sapien
4	921	64.4	337	6	Q95M34	Q95m34	equus caball
5	854	59.7	463	11	Q991C4	Q991c4	mus musculus
6	850	59.4	469	11	Q8R3V9	Q8r3v9	mus musculus
7	842	58.9	437	11	Q9R1A4	Q9r1a4	mus musculus
8	840	58.7	473	11	Q9D8L4	Q9d8l4	mus musculus
9	836	58.5	473	11	Q991Z5	Q991z5	mus musculus
10	830	58.0	468	11	Q99L31	Q99l31	mus musculus
11	806	56.4	474	11	Q8R3H6	Q8r3h6	mus musculus
12	801	56.0	473	11	Q91Z05	Q91z05	mus musculus
13	518	36.2	34350	4	Q9WZ42	Q9wz42	homo sapien
14	485	33.9	28926	4	Q10466	Q10466	homo sapien
15	485	33.9	29926	4	Q8WZB3	Q8wzb3	homo sapien
16	424	29.7	16215	5	Q9NFS3	Q9nfs3	drosohilla

```

Db 242 KSCDKTHTCCPCAPPELLGGPSVFLPPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNW 301
Qy 95 YVDGVEVHNAKTPREEQNSTYRVSVLTVLHQDLNGKEYCKVSNKALPAPIEKTIS 154
Db 302 YVDGVEVHNAKTPREEQNSTYRVSVLTVLHQDLNGKEYCKVSNKALPAPIEKTIS 361
Qy 155 KAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYFSPDSIAVWESNGQPENNYKTTPPV 214
Db 362 KAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYFSPDSIAVWESNGQPENNYKTTPPV 421
Qy 215 LDSGGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 264
Db 422 LDSGGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 471

RESULT 2
Q96PQ8 PRELIMINARY; PRT; 701 AA.
AC Q96PQ8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Factor VII active site mutant immunoconjugate.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21477448; PubMed=11593034;
RA Hu Z., Garen A.;
RT "Targeting tissue factor on tumor vascular endothelial cells and tumor
RT cells for immunotherapy in mouse models of prostatic cancer.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:12180-12185(2001).
DR EMBL; AF272774; AAK58686.1; -
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR000294; VitK_dep_GLA.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00594; gla; 1.
DR Pfam; PF00047; Ig; 2.
DR Pfam; PF00089; trypsin; 1.
DR SMART; SM00181; EGF; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS01187; EGF_CA; UNKNOWN_1.
DR PROSITE; PS00011; GLU_CARBOXYLATION; UNKNOWN_1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 701 AA; 77826 MW; 94AC6CEB42CC992F CRC64;

Query Match 87.5%; Score 1251; DB 4; Length 701;
Best Local Similarity 51.2%; Pred. No. 7.3e-41;
Matches 233; Conservative 8; Mismatches 9; Indels 205; Gaps 3;

Qy 7 YDKVAE----- 12
Db 255 FDKINRNLIAVLGEHDLSEHDEQSRRAQVIIPSTVPGTTHDIALRLHQPVWL 314
Qy 13 ----- 12
Db 315 TDHVVPLCLPTEFSERTAFVRESLVSGWQLDRGATALELMVLNVPRLMTQDCLQQS 374
Qy 13 ----- 12

```

```

Db 375 RKVGDSPNITPYMFCAGYSDGSKDSCAGDSGPHATHYRGWTYLTGIYSWGQGCATVGHF 434
Qy 13 ----KLKEAFDLKAFYDKVAEKLKEA-----PMDKTHTCPPCAP 49
Db 435 GYTVRSQYIEWL-----QKLMRSEPRGVLLRAPFGSAEPKSCDKTHTTCCPCAP 486
Qy 50 ELLGGPSVFLPPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTPR 109
Db 487 ELLGGPSVFLPPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTPR 546
Qy 110 EEOYNSTYRVSVLTVLHQDLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTL 169
Db 547 EEOYNSTYRVSVLTVLHQDLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTL 606
Qy 170 PSRDELTKNOVSLTCLVKGYFSPDSIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTV 229
Db 607 PSRDELTKNOVSLTCLVKGYFSPDSIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTV 666
Qy 230 DKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 264
Db 667 DKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 701

RESULT 3
Q8TC63 PRELIMINARY; PRT; 473 AA.
AC Q8TC63;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 52.0 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RC TISSUE=KIDNEY;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025985; AAH25985.1; -
KW Hypothetical protein.
SQ SEQUENCE 473 AA; 51986 MW; E29920B09BA369F5 CRC64;

Query Match 80.8%; Score 1155; DB 4; Length 473;
Best Local Similarity 48.8%; Pred. No. 1.5e-37;
Matches 219; Conservative 22; Mismatches 22; Indels 188; Gaps 5;

Qy 2 WLKA-----FYDKVAEKLKEAFDNL----- 21
Db 23 WLSRLQLQESGPGCLKPSVTLSTCTVSGDSVASSSYNGWVRQPPGKLEWIGTINFS 82
Qy 22 -----KAFYDKVAE----- 30
Db 83 GNMYYSPSLRSRVMTSADMSENSEFYKLDSTVTAADTAVYYCAAGHLYMGFCAHMGQGLV 142
Qy 31 -----KLKEAFMD----- 38
Db 143 SVSPASTKGPSVTFPLAPCSRSTSESTAALGCLVKDYFPEPVTVSWNSGALTSGVHTTTPAV 202
Qy 39 -----KTHT-----GPPCPAPELLG 53
Db 203 LQSSGLYSLSSVTVTPVSSSLGTYYTCNVDPKPSNTKVKRVSKEYGPPCSCPAPEFLG 262
Qy 54 GPSVFLPPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTPREEQY 113
Db 263 GPSVFLPPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTPREEQF 322
Qy 114 NSTYRVSVLTVLHQDLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRD 173
Db 323 NSTYRVSVLTVLHQDLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSQE 382
Qy 174 ELTKNOVSLTCLVKGYFSPDSIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSR 233

```



```

DE 383 EMKQVSLCLYKGFYPDIWESNGOPENNYKTPPVLDSDGSFFLYSLRTVDKSR 442
QY 234 WOQNVFSCVMHEALHNHYTKSLSPGK 264
DB 443 WOQNVFSCVMHEALHNHYTKSLSLGK 473

RESULT 4
Q95M34 PRELIMINARY; PRT; 337 AA.
AC Q95M34;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Immunoglobulin gamma 1 heavy chain constant region
DE (Fragment).
GN IGHC1.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RA Wagner B.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98383416; PubMed=9717671;
RA Wagner B., Overesch G., Sheoran A., Holmes M., Richards C.,
RA Leibold W., Radbruch A.;
RT "Organization of the equine immunoglobulin heavy chain constant region
RT genes. III. Alignment of c-mu, c-gamma, c-epsilon and c-alpha genes.";
RL Immunobiology 199:105-119(1998).
DR EMBL; AJ300675; CAC44624.1;
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 2.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.
FT NON_TER
FT 1
SQ SEQUENCE 337 AA; 37438 MW; A60BF2B01DEFD1F6 CRC64;

Query Match 64.4%; Score 921; DB 6; Length 337;
Best Local Similarity 60.8%; Pred. No. 1.4e-28;
Matches 169; Conservative 36; Mismatches 42; Indels 31; Gaps 4;

QY 2 WLKAFY-----DKVAEKLKEAFDNLKAFYDKVAEKLKEAFMDKTHTCPCPA 48
DB 76 WTSETYICNVVHAASNFKVDKRIEIPDN-----HOKVCDMSK-----CPKCPA 119
QY 49 PELLGGPSVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKP 108
DB 120 PELLGGPSVFLFPNPKDLMITPTETVCVVVDVSDQENPDVKNWYMDGVEVETATRP 179
QY 109 REEQNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKAKGQPREPVYTL 168
DB 180 KEEQNFSTYRVSVLRIQHODWLNKGEYKCKVSNKALPAPIEKTISKAKGQPREPVYTL 239
QY 169 PPSRDELTKNOVSLTCLVKGFPYSDIAVEWESNGQP--ENNYKTTTPVLDSDGSFFLYSK 226
DB 240 APHPDELKSKSVTCLVKDIPPEINIEWQSNQGPPELTKYSTTQAQDSDGSYFLYSK 299
QY 227 LTVDKSRWQOQNVFSCVMHEALHNHYTKSLSPGK 264
DB 300 LSVDRNRWQOQNTTTCGVWHEALHNHYTKVSNKPK 337

RESULT 5
Q99LC4 PRELIMINARY; PRT; 463 AA.
AC Q99LC4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

```

```

DE Similar to RIKEN CDNA 1810060009 gene.
GN IGH-4.
OC Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003435; AAH03435.1;
DR HSSP; P01842; 7FAB.
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00409; Ig; 2.
DR SMART; SM00407; IGcl; 2.
DR SMART; SM00406; IGv; 1.
DR SMART; SM00410; IG_like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 463 AA; 51007 MW; EAA674C6BBC30783 CRC64;

Query Match 59.7%; Score 854; DB 11; Length 463;
Best Local Similarity 37.0%; Pred. No. 2.1e-25;
Matches 155; Conservative 50; Mismatches 48; Indels 166; Gaps 8;

QY 2 WLKAFYDKVAEKLKEAFDNLKAFY-----DKVAEKLK-----EAFMD----- 38
DB 55 WVK-----QRTQGLEWGEIYPGSGNTYSEKFKGKATLTDDKSSYATMHLSSLTS 107
QY 39 -----HT-----KT----- 40
DB 108 EDSAVFYFCARSSYYSDLFAYWGQGLTVTSAKTTTPSVYPLAPGSAQAQNSMVTGLCL 167
QY 41 -----HT----- 42
DB 168 VKGFPEPVTVWNSSGLSSGVHFFPAVLQSDLYLTLSSTVTPSPSTPSEVTCNVAHPA 227
QY 43 -----CPP--CPAPELGGPSVFLFPKPKDLMISRTPEVTCVVVDVSH 85
DB 228 SSTRVDKKIVPRDCGCKPCICTVPEV---SSVFIFPPKPKDVLITLTPKTCVVVDISK 284
QY 86 EDPEKFNWYVDGVEVHNAKTKPREEQYNSTYRVSVLTVLHODWLNKGEYKCKVSNKAL 145
DB 285 DDPEVQFSWFVDYEVHTAQTPREQFNSTFRSVSELPIMHODWLNKGEYKCKVNSAAF 344
QY 146 PAPIEKTISKAKGQPREPVYTLPPSRDELTKNOVSLTCLVKGFPYSDIAVEWESNGOPE 205
DB 345 PAPIEKTISKAKGQPAQVYIIPPKQMAKDKVSLFCMTDFFPEDITVWQWNGQPA 404
QY 206 NNYKTTTPVLDSDGSFFLYSKLTVDKSRWQOQNVFSCVMHEALHNHYTKSLSPGK 264
DB 405 ENYKNTQPIQNDTSGSYFLYSLKLVNOKSNWEAGNFTCSVLHEGLHNHHHTSLSHSPGK 463

RESULT 6
Q8R3V9 PRELIMINARY; PRT; 469 AA.
AC Q8R3V9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 52.0 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

```

RA Strausberg R.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC024405; AAH24405.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 469 AA: 51976 MW: 534793F15D05457 CRC64;

Query Match 59.4%; Score 850; DB 11; Length 469;
 Best Local Similarity 36.7%; Pred. No. 3.2e-25;
 Matches 151; Conservative 47; Mismatches 48; Indels 165; Gaps 5;

QY 16 EAPDWL----- 21
 : : : :
 Db 62 KALEWLFIRNKANGYTEYSASVKGRTISRDSQSILYQMNALRAEDSATYYCARDR 121
 : : : :
 QY 22 -----KAFYDK- 27
 : : : :
 Db 122 RSYYYSGTSFAYWGGTGLTVTSAAKTTPPSVYPLAPGSAATNSMVTGLGVLKGYFFPP 181
 : : : :
 QY 28 -----VAEKLKEAFMDKT 40
 : : : :
 Db 182 VVTWNSGSLSSGVHTFPVQLSDLYTLSSSVTVPSSTWPSQTVTCNVAPASSTKVDKK 241
 : : : :
 QY 41 HT-----CPP--CPAPELLGGPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFN 93
 : : : :
 Db 242 IVPDGGCKPCICTVPEV---SSVFIPPPKPKDVLITLTPKVTCTVVDVSKDDPEVQFS 298
 : : : :
 QY 94 WVVVDGVEVHNATKPREOYNSTYRVSVTLVLDHDLNGKCYKCKVSNKALPAPIEKT 153
 : : : :
 Db 299 WFDVDDVEVHTAQTPREEQFNSTFRSVSELPIMHQDLNGKFKCRVNSAAPAPIEKT 358
 : : : :
 QY 154 SKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYFSPDIAVEWESNGQPENNYKTTTP 213
 : : : :
 Db 359 SKTKGRPKAPQVYTIPTPPKEQMAKDKVSLTCLTMITDFEPDITVEMQWNGQPAENYKNTQ 418
 : : : :
 QY 214 VLDSGDSFPLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKQSLSPGK 264
 : : : :
 Db 419 IMDTDSGYFYSLKLVNOKSNWEAGNTFTCSVLHGLHNNHTEKLSLSPGK 469
 : : : :

RESULT 7

Q9R1A4
 ID Q9R1A4 PRELIMINARY; PRT: 437 AA.
 AC Q9R1A4;
 DT 01-MAY-2000 (TremBLrel. 13, Created)
 DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
 DE Gamal heavy chain of Mab7 (Fragment).
 GN IGH-4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wilde K.G., Yu X., Ekramoddoullah A.K.M., Misra S.;
 RT "Cloning of cDNAs encoding for anti-white pine blister rust monoclonal
 RT antibody (Mab 7, its light and heavy chains) and construction of a
 RT single chain antibody (scFv).";
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF152372; AAD40243.1; -
 DR HSP; P01842; 7FAB.
 DR MGD; MGI:96446; Igh-4.
 DR InterPro; IPR003600; Ig_like.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003596; Ig_v.
 DR Pfam; PF00047; Ig_4.
 DR SMART; SM00406; Igv; 1.
 DR SMART; SM00410; IG_like; 2.
 DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
 FT NON_TER 1
 FT NON_TER 437
 SQ SEQUENCE 437 AA: 48142 MW: 5C3AY7BB3E7D697C CRC64;

Query Match 58.9%; Score 842; DB 11; Length 437;
 Best Local Similarity 36.3%; Pred. No. 5.3e-25;
 Matches 150; Conservative 51; Mismatches 52; Indels 160; Gaps 6;
 QY 2 WUKAFYDKVAEKLKEAFMDKAF-----YDKVAEKLKEAFMDK----- 39
 : : : : : : : :
 Db 35 WVRQTPKEK-----RLEWVSFSGGIYITDYSVKGRFTIYKDKDRNLSLQMSLSRSE 87
 : : : : : : : :
 QY 40 ----- 39
 : : : : : : : :
 Db 88 DTAMYICARGDSYAWGPGTGLTVTSAAKTTPPSVYPLAPGSAATNSMVTGLGVLKGYPP 147
 : : : : : : : :
 QY 40 -----THT----- 42
 : : : : : : : :
 Db 148 EPVTVTWSGSLSSGVHTFPVQLSDLYTLSSSVTVPSSTWPSQTVTCNVAPASSTKVD 207
 : : : : : : : :
 QY 43 -----CPP--CPAPELLGGPSVFLPPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 91
 : : : : : : : :
 Db 208 KIVPRDGGCKPCICTVPEV---SSVFIPPPKPKDVLITLTPKVTCTVVDVSKDDPEVQ 264
 : : : : : : : :
 QY 92 FNNYVDGVEVHNATKPREOYNSTYRVSVTLVLDHDLNGKCYKCKVSNKALPAPIEKT 151
 : : : : : : : :
 Db 265 FSNFVDDVEVHTAQTPREEQFNSTFRSVSELPIMHQDLNGKFKCRVNSAAPAPIEKT 324
 : : : : : : : :
 QY 152 TISKAGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYFSPDIAVEWESNGQPENNYKTT 211
 : : : : : : : :
 Db 325 TISKTKGRPKAPQVYTIPTPPKEQMAKDKVSLTCLTMITDFEPDITVEMQWNGQPAENYKNT 384
 : : : : : : : :
 QY 212 PVLDSGDSFPLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKQSLSPGK 264
 : : : : : : : :
 Db 385 QPIMDTDSGYFYSLKLVNOKSNWEAGNTFTCSVLHGLHNNHTEKLSLSPGK 437
 : : : : : : : :
 RESULT 8
 Q9D8L4
 ID Q9D8L4 PRELIMINARY; PRT: 473 AA.
 AC Q9D8L4;
 DT 01-JUN-2001 (TremBLrel. 17, Created)
 DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
 DE 1810060009Rik protein.
 GN IGH-1 OR 1810060009Rik.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Alizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojbori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikola I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boiffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kaniya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 DR EMBL: AK007918; BAB25349.1; -
 DR HSP; P01842; 7FAB.
 DR MGD; MGI:96443; Igh-1.

```

DR InterPro: IPR003599; Ig.
DR SMART: SM00409; IG; 2.
DR SMART: SM00407; IGcl; 3.
DR SMART: SM00406; IGV; 1.
DR SMART: SM00410; IG_like; 1.
DR SMART: SM00407; IGcl; 3.
DR SMART: SM00406; IGV; 1.
DR SMART: SM00410; IG_like; 1.
DR SMART: SM00409; IG; 2.
DR SMART: SM00407; IGcl; 3.
DR SMART: SM00406; IGV; 1.
DR SMART: SM00410; IG_like; 1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 473 AA; 51699 MW; 9DED57A514475FBB CRC64;

Query Match      58.7%; Score 840; DB 11; Length 473;
Best Local Similarity 36.4%; Pred. No. 8.3e-25;
Matches 157; Conservative 46; Mismatches 50; Indels 178; Gaps 7;

QY 1 DWLK-----AFYDKVAEKLKEA----- 17
Db 54 NWYKQRPQGLEWIGKIGPGSGSTYIN---EKFKGKATLTADKSSSTAYMQLSSTSEDS 110
QY 18 -----FDW-----LKAF 24
Db 111 AVYFCARSGYDWDFAWYWGQGLTVTSAAKTAPSVYPLAPVCGGTGSSVTGLGLVKGY 170
QY 25 Y----- 25
Db 171 FPEPVTLTWNSGSLSGVHTFPALLQSGLYTLSSSVTVTSNTWPSQTIITCNVAHPASSTK 230
QY 26 -DKVAE-----KLKEAFMDKTHTCPAPPELLGGPSVFLFPKPKDITLMISRT 73
Db 231 VDKKIEPRVITONPCPPPLKE-----CPPCAAPDLLGGPSVFIIPPKIKDVLMSLS 282
QY 74 PEYTCVVVDYSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNG 133
Db 283 PMTCVVVDVSDPDQVQISWPNVNVETVHTAQTQTHREDYNSLTIRVVSALPIQHODWMSG 342
QY 134 KEYKCKYSNKKALPAPIETKISKAGQOPREPQVYTLPPSRDELTKNOVSLTCLVKGFPYPSD 193
Db 343 KEFKCKVNNRALSPIETKISKRGVPRAPQVYVLPPEAEETKKEESLTCMITGFLPAE 402
QY 194 IAVENESNGQPNKYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFCSMHEALHNNHY 253
Db 403 IAVDWTSTNGRTEQYKNTATVLDSDGSYFMYSKLRVQKSTWGRGSLFACSWHEGLHNL 462
QY 254 TQKSLSLSPCK 264
Db 463 TTKTISRSLGK 473

RESULT 9
Q99L25 ID Q99L25 PRELIMINARY; PRT; 473 AA.
AC Q99L25;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to RIKEN CDNA 1810060009 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003888; AA003888.1;
DR HSSP; P01842; 7FAB.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003597; Ig_cl.
DR InterPro: IPR003600; Ig_like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.

QY 254 TQKSLSLSPCK 264
Db 463 TTKTISRSLGK 473

RESULT 10
Q99L31 ID Q99L31 PRELIMINARY; PRT; 468 AA.
AC Q99L31;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to RIKEN CDNA 1810060009 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003878; AA003878.1;
DR HSSP; P01842; 7FAB.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003597; Ig_cl.
DR InterPro: IPR003600; Ig_like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGcl; 3.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_like; 1.

```

```

DR Pfam; PF00047; Ig; 3.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGcl; 3.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_like; 1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 473 AA; 52449 MW; BE9889B7986DA155 CRC64;

Query Match      58.5%; Score 836; DB 11; Length 473;
Best Local Similarity 37.6%; Pred. No. 1.2e-24;
Matches 160; Conservative 38; Mismatches 58; Indels 170; Gaps 7;

QY 2 WLKAFYDKVAEKLKEAFDWLKAFF-----DKVAEKLK-----EAFM----- 37
Db 55 WVK-----QRPEQGLEWIGYIYPRDGSSTKYNEKPKGKATLTADKSSSTAYMQLSLS 107
QY 38 ----- 40
Db 108 EDSAVCFCSRGSYIYGYLYFYDWYQGQTTITVSSAKTTAPSVYPLAPVCGDTTGSVT 167
QY 41 -----HT----- 42
Db 168 LGLVKGYFPEPVTLTWNSGSLSGVHTFPVQLQSDLYTLSSSVTVTSWPSQITCNV 227
QY 43 -----CPP--CPAPELLGGPSVFLFPKPKDITLMISRTPEVTC 78
Db 228 AHPASSTKVDKKTIEPRGPTIKPCPCPCAPNLLGGPSVFIIPPKIKDVLMSLSPMVTC 287
QY 79 VYVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKC 138
Db 288 VYVDVSEDDPDQVQISWPNVNVETVHTAQTQTHREDYNSLTIRVVSALPIQHODWMSGKEFKC 347
QY 139 KVSNNKALPAPIETKISKAGQOPREPQVYTLPPSRDELTKNOVSLTCLVKGFPYPSDIAVEW 198
Db 348 KVNKKALPAPIETKISKAGVSRAPQVYVLPPEEETKKQVTLTCMTDFMPEDIIVEW 407
QY 199 ESNQGNPNKYKTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFCSMHEALHNNHYTKSL 258
Db 408 TNGKTELKYNKTEPVLDSGSYFMYSKLRVQKSTWGRGSLFACSWHEGLHNNHTKSF 467
QY 259 SLSPGK 264
Db 468 SRTPGK 473

RESULT 10
Q99L31 ID Q99L31 PRELIMINARY; PRT; 468 AA.
AC Q99L31;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to RIKEN CDNA 1810060009 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003878; AA003878.1;
DR HSSP; P01842; 7FAB.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003597; Ig_cl.
DR InterPro: IPR003600; Ig_like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGcl; 3.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_like; 1.

```

```
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 468 AA; 51661 MW; 96352328B3332ADB CRC64;

Query Match
Best Local Similarity 58.0%; Score 830; DB 11; Length 468;
Matches 158; Conservative 38; Mismatches 55; Indels 175; Gaps 7;

QY 2 WLKAFYDKVAEKLKEAFDMLKAFYD-----KVAEKLKEAFMDK----- 39
Db 55 WK-----QRPEQGLEWI-GWIDPEDGETKYAPK-----FQDKATITADTSSNTAYLQL 102
QY 40 -----THT----- 39
Db 103 SSLTSEDTAIYCARNLGYGYDYGQGTITVSSAKTTAPSVYPLAPVCGDITGSSVT 162
QY 40 -----THT----- 42
Db 163 LGLLVKGYPPEVTLTWNSGSLSSGVHTFPVQLSDLYLTSSVTVTSWPSSQITCNV 222
QY 43 -----CPP--CPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTC 78
Db 223 AHPASSTKVDDKLEPRGPTIKPCPCCKCPAPNLGGPSVFIIPPNIKIDVLMISLSPWTC 282
QY 79 VVVDVSHEDPEVKFNWYDGVVEVHNNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYK 138
Db 283 VVVDVSEDDPDVQISWNNVEVLTATQTHREDYNSTLRVVSALPFIHQDWMNGKEYK 342
QY 139 KYSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEW 198
Db 343 KYNKALPAPIERTISKAGSVRAQVYVLPPEEEMTKKQVTLTCMVYDFMPEDIDYEW 402
QY 199 ESNQGPENNYKTPPPVLDSDGSFFLYSLKTVDSKRWQGNVFCSCVMHEALHNYTKSL 258
Db 403 TNGKTELAKNTEPVLDSGDSGYFMYSLRVEKKNVENSYSVCSVWVHGLNHHHTKSF 462
QY 259 SLSPGK 264
Db 463 SRTPGK 468

RESULT 11
Q8R3H6 PRELIMINARY; PRT; 474 AA.
AC Q8R3H6;
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DE Hypothetical 51.7 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025447; AAH25447.1;
DR MGD; MGI:2144967; A0044919.
DR InterPro; IPR003345; CytC_heme_bind.
DR Pfam; PF00047; Ig_3.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 474 AA; 51748 MW; 8608B57C6CD2874A CRC64;

Query Match
Best Local Similarity 56.4%; Score 806; DB 11; Length 474;
Matches 151; Conservative 44; Mismatches 59; Indels 177; Gaps 6;

QY 1 DWLKA----- 5
Db 54 NNWKRPRGKLEWIGRIFFGCGDTHYSGKQAKLTADKSSVTAFLQTSITSEDSAVY 113
QY 6 -----FYDKVAEKLKEAFD----- 20
Db 114 FCARDSYGDYFD-----DNQOGATVTVSSAKTTPPSVYPLAPCGDITGSSVT 163
QY 21 ---LKAFYDK-----VA 29
```

```
Db 164 GCLVKGYPPEVTVTWNSGSLSSVHTFPALLOSLGLYTMSSSVTVPSSTWPSQTVTCSVA 223
QY 30 EKLKEAFMDK-----THTCPP-----CPAPPELLGGPSVFLFPPKPKDTLMISRT 73
Db 224 HPASSTTVDDKLEPSGPISTINPCPCCKECHKCPAPNLEGPSVFIIPPNIKIDVLMISLT 283
QY 74 PEVTCVVVDVSHEDPEVKFNWYDGVVEVHNNAKTPREEQYNSTYRVVSVLTVLHQDWLNG 133
Db 284 PKVTCVVVDVSEDDPDVQISWNNVEVLTATQTHREDYNSTLRVVSALPFIHQDWMNGSG 343
QY 134 KEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSD 193
Db 344 KEFKCKVNNKLPSPERTISKIGLVRAPQVYVILPPPAEQLSRKDVSLTCLVGFENPGD 403
QY 194 IAVVESNGQPPENNYKTPPPVLDSDGSFFLYSLKTVDSKRWQGNVFCSCVMHEALHNY 253
Db 404 TSVETNSGTENTKDTAPVLDSDGSFYIYSLKDIKTSKWEKTDTSFSCNVRHEGLKNYY 463
QY 254 TOKSLSLSPGK 264
Db 464 LKKTISRSPGK 474

RESULT 12
Q91Z05 PRELIMINARY; PRT; 473 AA.
AC Q91Z05;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 51.9 kDa protein.
GN AU044919.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010327; AAH10327.1;
DR MGD; MGI:2144967; A0044919.
DR InterPro; IPR003345; CytC_heme_bind.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_3.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 473 AA; 51946 MW; CF625F008932AF12 CRC64;

Query Match
Best Local Similarity 56.0%; Score 801; DB 11; Length 473;
Matches 153; Conservative 44; Mismatches 66; Indels 156; Gaps 6;

QY 2 WLKAFYDK-----VAEKLKEAFD----- 19
Db 55 WVRQAPKGLWVAYVINGSTTIYVADTVKGRFTISRDNAKNTFLQMTSLRSEDATAMY 114
QY 20 -----WL-----KAFYDK--- 27
Db 115 CARELWLRRIYWGQGTITVSSAKTTPPSVYPLAPGCGDITGSSVTGLVKGFYPSV 174
QY 28 -----VAEKLKEAFMDK--- 39
Db 175 TVTWNSGSLSSVHTFPALLOSLGLYTMSSSVTVPSSTWPSQTVTCSVAHPASSTWPKDL 234
QY 40 -----THTCPP-----CPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSH 85
Db 235 EPSGPISITINPCPCCKECHKCPAPNLEGPSVFIIPPNIKIDVLMISLTPKTVVVDVSE 294
QY 86 EDPEVKFNWYDGVVEVHNNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKAL 145
```


Db 2549 NVFVELSHSGIDVLWNFKDKETKPPSKYKIEAHGKIYKLTVLNMMKDEGKYFYAGE 2608
Qy 10 ----- 9
Db 2609 NITSGKLTVAGAIKPLTDQTVAESQAEVPECEVANPDSKGWLRDGHPLPTNNIRSE 2668
Qy 10 ----- 9
Db 2669 SDGKRRLLIAATKLDIDIGEYTKVATSKTSAKLKVEAVIKKTLKNTLTVTQDAVTV 2728
Qy 10 ----- 9
Db 2729 ELTHPNVGVOMKNGVVLESNEKYAISVKGTIYSLRIKNCIIVDES YGFRGLRGASA 2788
Qy 10 ----- 9
Db 2789 RLHVETVKIIKKPRDVTALENATVAFVSVDHVPVFWFHKNVEIKPSDKHRLVSEKVV 2848
Qy 10 ----- 9
Db 2849 HKMLONISPDAGEYTA VGOLECKAKLFVETLHITKTMKNIEVPETKTASFCEVSHF 2908
Qy 10 ----- 9
Db 2909 NVPSMWLKNVIEIEMSEKFI VOGKHLHOLIIMNTSTEDSAEYTFVCGNDQVSATLAVTP 2968
Qy 10 ----- 9
Db 2969 IMITSMKLDINAEKDTITFEVTNVEGISYKWLKNGVEIKSTDKCQMRTKKLTHSLNIR 3028
Qy 10 ----- 9
Db 3029 NVHFGDAADYTFVAGKATSTATLYVEARHIEFRKHIDIKVLEKKRAMFECEVSEPDITV 3088
Qy 10 ----- 9
Db 3089 QMVKDDQELQITDRIKIQEKYVHRLLIPSTRMSDAGKYTVVAGNVSTAKLFVEGRDVR 3148
Qy 10 ----- 9
Db 3149 IRSIKKEVQVIEKQAVVEFEVNEDDVDHMYKDGIEINFQVQERHKYVVERIHRMFIS 3208
Qy 10 ----- 9
Db 3209 ETRQSDAGEYTFVAGNRSSVTVLYVNAPEPPQVLQELQVTVQSGKPARFCAVISGRPOP 3268
Qy 10 ----- 9
Db 3269 KISWYKEBQLLSTGFKCKFLHDGQEYTLILLIEAFPEDAAYTCEAKNDYGVATTSSLSV 3328
Qy 10 ----- 9
Db 3329 EVPEVSPDQEMPVYPPIAITPLQDVTVSEGPQARFQCRVSGTDLKVSWSKDKKIKPSR 3388
Qy 10 ----- 9
Db 3389 FFRMTQPDYQLEIAEAYPEDEGTYTFVASNAGVQVSSSTANLSLEAPESILHERIQEI 3448
Qy 10 ----- 9
Db 3449 EMEMKEFSSFLSABEEGLHSAELQSKINETLELLSESPVYSTKFDSEKEGTGPIFIKE 3508
Qy 10 ----- 9
Db 3509 VSNADISMGDVATLSVTVIGIPKPIQWFPNGVLLTPSADYKFFVFDGDDHSLIILFTKLE 3568
Qy 10 ----- 9
Db 3569 DEGEYTCMASNDYKGTICSAVYLKINSKGECHKDTETESAVAKSLEKLGCPPHFLKELK 3628
Qy 10 ----- 9
Db 3629 PIRCAQGLPAIFETVTVGEPAPTVPFKENKOLCTSVYTIHNPNGSGTFIVNDPQRED 3688

Qy 10 ----- 9
Db 3689 SGLYICRAENMLGESTCAEALLVLLLEDTDMDTTPCKAKSTPEAPEDFPQTPLKGPAGEAL 3748
Qy 10 ----- 9
Db 3749 DSEQIATFVKDTILKAALITEENQOLSYEHIAKANELSSQLPLGAQELQSLQEQDKLTP 3808
Qy 10 ----- 9
Db 3809 ESTREFLCINGSIHQPLUKBSPNLQLOIVOSQKTFSEKILMPEEPETQAVLSOTEKIF 3868
Qy 10 ----- 12
Db 3869 PSAMSIEOINSITVEPLKTLAEPEGNYQSSIEPPHMSYLTSAEEVLSPEKTVSDTN 3928
Qy 13 ----- 12
Db 3929 REQVTLQKQEAQSALILSQAEGHVESLQSPDVMISQVNYEPLVPSEHSCTEGGKILI 3988
Qy 13 ----- 12
Db 3989 ESAMPLENAGODSAVRIEEGKSLRFPPLALEBKQVLLKEEHSNDNVMPDQIESKREPA 4048
Qy 13 ----- 12
Db 4049 IKKVOEVQGRDLLSKESLLSGIPBEORLNLKIQICRALQAASVASEPGLFSEWLRNIEKV 4108
Qy 13 ----- 12
Db 4109 EVEAVNITQEPHRIMCYLVTSAKSVTEVTIIIEDVDPQMANLKMELRDALCAIYEI 4168
Qy 13 ----- 12
Db 4169 DILTAEGPRIQOAGKTSLOEEMDSFGSKVEPITEPEVESKYLSISPEVSYFNQSVRK 4228
Qy 13 ----- 17
Db 4229 YLDATPVTKGVASAVVDEKQDESILKPSEKESESSESGTEEVATVKIQEAGGFIKEDG 4288
Qy 18 ----- 17
Db 4289 PMIHPLVDTVSEBGDIVHLTTSITNAKEVNNYFENKLVPSDEKFKCLQDQNTVTLIDK 4348
Qy 18 ----- 17
Db 4349 VNTEHOGYVCEALNDSGKTATSAKLTVVKRAAPVIRKRIEPLVALGHLAKFTCEIQS 4408
Qy 18 ----- 23
Db 4409 APNVRFQWFKAGRIEYSDKCSIRSSKYISSLEILRTQVDCGEYTCASNEYGSVSCA 4468
Qy 24 ----- 23
Db 4469 TLTVTEAVPPTFLSRPKSLTTFVGKAAKFICTVGTVPVETIWKDGAALSPSPNMKISD 4528
Qy 24 ----- 23
Db 4529 AENKHILELSNLTIQDRGVYSCASNKFGADICQAEILIIDKPHFIKELEPVQSAINKKV 4588
Qy 24 ----- 29
Db 4589 HLEQVDEDRKVTVTWKDGOKLPPGKDYKICFEDKATLEIPLAKLSDSYVCTASNE 4648
Qy 30 ----- 29
Db 4649 AGSSSCSATVTVREPPSFVKVVDPSYMLPGESARLHCKLKGPSVIOVTFWFKNNKELSE 4708
Qy 30 ----- 29
Db 4709 NTRMYFVNSEALIDITDKVEDSGSYSCAENVDSGSDCSTEIVIKEPPSFIKTLEPAD 4768

QY 30 ----- 29
Db 4769 IVRGTNALLQCEYSGTGPPEISWFKDKQIRSSKKYRLFQKSLVCLLEIFSPNSADVGEY 4828
QY 30 ----- 29
Db 4829 ECVVANEVKGCGMATHLLKEPPTFKVKVDLLIALGGQVTLQAAVRGSEPISVTWKMGQ 4888
QY 30 ----- 29
Db 4889 EVIREDGKIKMSFNGVAVLLIIPDVQISFGCKYTCLAEENAGSQTSVGELIVREPAKIIIE 4948
QY 30 ----- 29
Db 4949 RAEIIQVAGDPATLEYTVAGTPELPKWKYKDGRLPLVASKKYRISFKNNVAQLKFYSAEL 5008
QY 30 ----- 29
Db 5009 HDSGQYTFEISNEVGSSCETTFVLDRDIAPFTFKPLRNVDVSVVNGTCRLDCKIAGSLP 5068
QY 30 ----- 29
Db 5069 MRVSWFKDGKEIAASDRYIAFVEGTASLEIIRVDMNDAGNFTCRATNSVGKDSGALI 5128
QY 30 ----- 29
Db 5129 VQPPSFVTKPGSKDVLPGSAVCLKSTFGGSTPLTIIRWFKGNKELVSGGSCYITKEALES 5188
QY 30 ----- 29
Db 5189 SLELYLVKTSDSGTCTCKVSNVAGGVECSANLFVKEPATFVEKLEPSQLLKGDATQLAC 5248
QY 30 ----- 29
Db 5249 KVTGTPPIKITWFANDREIKESKHRMSFVESTAVLRLTDVGIEDSGEYACEAONEAGSD 5308
QY 30 ----- 29
Db 5309 HCSSIVIVKESPYTFKEFKPIEVKEYDVMLAEVAGTPPEITWFKDNTILRSGRKYKT 5368
QY 30 ----- 29
Db 5369 FIODHLVSLQILFVAADAGEYOCRVTVNEVGSSIC SARVTLREPPSPFIKKIESTSLRG 5428
QY 30 ----- 29
Db 5429 TAAQATLKGSPLITYTWLKDSDEITEDDNIRMTFENNVAASLYLSGIEVKHDKGYVCQAK 5488
QY 30 ----- 29
Db 5489 NDAGIORCSALLSVKEPATITEEAVSIDVTQGD PATLQVNFSGTKETAKWFKDGOELTL 5548
QY 30 ----- 29
Db 5549 GSKYKISVTDVTSILKIISTEKKDGEYTFEVQNDVGRSSCKARINVLDLIIPPSFTKKL 5608
QY 30 ----- 39
Db 5609 KKMDSIKGSFIDLECI VAGSHPI SIQWFKDDQEI SASEKYKFSFHDNTAFLEISQLEBTD 5668
QY 40 --THTC-----PP----- 45
Db 5669 SGTYTCSATNKGHNQCSGHLTVKEPPYFVEKPKSQSDVNPNTRVOLKALVGGTAPMTIKW 5728
QY 46 -----C----- 46
Db 5729 FKONKELHSGAARSWKDDTSTSLFPAKATDSGTICYCOLSNDVGTATSKATL FVKEPP 5788
QY 47 ----PAP----- 49
Db 5789 QFIKKPSVLVLRNGOSTTFECQITGTPKIRVSWYLDGNEITAIQKHGISFIDGLATFOI 5848
QY 50 ----- 49

Db 5849 SGARVENSGTYVCEARNDA GTASCSIELVKVEPPTFIRELKPVVEVVKYSDVELECEVTGT 5908
QY 50 ----- 49
Db 5909 PPFEVTLKNNREIRSSKKYTLDRVSVFNLHITKCDPSDTGEYOCIVSNEGSGSCSTR 5968
QY 50 ----- 49
Db 5969 VALKEPPSPFIKKIENTTIVLKSSATFQSTVAGSPPI SITWLKDDQIILDEDDNVVISFVDS 6028
QY 50 ----- 49
Db 6029 VATLQIRSVDNHSGRYTQCAKNESGVERCYAFLLVQEPQIVAKAKSVDTVTEKDPMTLE 6088
QY 50 ----- 49
Db 6089 CVVAGTPELVKWLKDGKQIVPSRYFSMSPENNVA SFRIOSVMKQDSQGYTFKVENDFGS 6148
QY 50 ----- 49
Db 6149 SSCDAYLRVLDQNI PPSTFKLTKMDKVLGSSIHMECKVSGSLPISAQWFKDGKEISTSA 6208
QY 50 ----- 49
Db 6209 KYRLVCHERSVSVLENNLEDTANYTCVSNVAGDDACSGILTVKEPPSFLVKPGRQQA 6268
QY 50 ----- 59
Db 6269 IPDSTVEFKAILKGT PPFKIKWFKDDVELVSGPKCFITGLEGSTFSLNLYSVDASKTQGYT 6328
QY 60 ----- 59
Db 6329 CHVTNDVGSCTTMLLVTEPPKFVKLEASKIVKAGDSSRLECKIAGSPEIRVVWFRNE 6388
QY 60 ----- 59
Db 6389 HELPASDKYRMTFIDSVAVIQMNNLSTEDSGDFICEAQNPAGSTSCSTKVIVKEPPVFSS 6448
QY 60 FPP----- 62
Db 6449 FPIVETLKNAEVSLECELSGT PPFVYVWYKDKRQLRSSKKYKIASKNFHTSIHLNVD 6508
QY 63 ----- 62
Db 6509 SDIGEXHCKAQNEVGSDTCVCTVKLKEPPRFVSKLNSLITVVAGEPAELQASIEGAQPIFV 6568
QY 63 ---KPKD-----TLM-- 69
Db 6569 QWLKEKEEVIRESNIRITFVENVATLQFAKAPANAGKYICQIKNDGGMEENATLMVL 6628
QY 70 ----- 69
Db 6629 EPAVIVEKAGPMTVTGETCTLECKVAGTPELSVWYKDGKLLTSSQKHKFSFYNKISSL 6688
QY 70 ----- 69
Db 6689 RILSVERQDAGTYTFQVQNNVGKSSCTAVVDVSDRAVPSPFTRLKNTGGVLGASCILEC 6748
QY 70 ----- 69
Db 6749 KVAGSSPISVAFHEKTI VSGAKYQTTFS DNVTCTQLNSLDSDMGNYTCVAANVAGSD 6808
QY 70 ----- 69
Db 6809 ECRAVLTVQEPSPFVKEPEPEPLVLPKNVTFTSVIRGTPPFKVNFRGARELVKGDRCNI 6868
QY 70 -----ISRTPEVTCV----- 80
Db 6869 YFEDTVAELFLNIDISQSGEYTCVSNNAQASCTTFLFVKPEAPAFKRLSDHSVEPGK 6928
QY 81 ----- 80

Db 6929 SIILESTYTGTLPISVTWKKGDNITTEKCNIVTTEKTCILEILNSTKRDAGQYSCIEI 6988
QY 81 -----FNNYVDGV----- 99
: || |||
Db 6989 NEAGRDVCGALYSTLEPPYFVTELEPLEAAGDSVSLQOVAGTPEITVSWYKGDTKLRP 7048
QY 81 -----FNNYVDGV----- 99
: || |||
Db 7049 TPEYRTYFTNNVATLVFNKNVINDSGEYTKAENSIGTASSKTVPRIQERQLPPSFARQL 7108
QY 81 -----FNNYVDGV----- 99
: || |||
Db 7109 KDIEQTGLVPLTLCRLNGSAPIQVCWYRDGVLLRDRDENLOTFSFVNDVATLKILQTDLSH 7168
QY 86 E-----FNNYVDGV----- 99
: || |||
Db 7169 SQYSCSASNPGLTASSARLTAREPKKSPFFDIKPVSIDVJAGESADFECHVTGAQPMR 7228
QY 87 -----FNNYVDGV----- 99
: || |||
Db 7229 ITWSKDNKEIRPGGNYTITCVGNTPHLRILKVGKDSGQYTCQATNDVYKDMCSAQLSVK 7288
QY 87 -----FNNYVDGV----- 99
: || |||
Db 7289 EPPKFVKLEASKVAKQGESIQLECKISGSPEIKVSWFRNDSSELHESWKYNMFSINSVAL 7348
QY 92 -----FNNYVDGV----- 99
: || |||
Db 7349 LLINEASAEADSGDYICEAHNGVGDASCSTALTATVKAAPPVFTQKPSVPGALKGSDVILQCEI 7408
QY 92 -----FNNYVDGV----- 99
: || |||
Db 7409 SCTPPEVWVKDRKQVRNKKFKITSKHFDLSHLINLEASDVGEYHCKATNEVGSOTC 7468
QY 92 -----FNNYVDGV----- 99
: || |||
Db 7469 SCVAFKEPPFVVKLSDTSTLIGDAVELRAIVEGFQPIVSVWLKDRGEVIRESENTRIS 7528
QY 92 -----FNNYVDGV----- 99
: || |||
Db 7529 FIDNIATLQGSPEASNSGKYICQKNDAGMRECSAVLTVLEPARIIEKPEPMTVTGNP 7588
QY 92 -----FNNYVDGV----- 99
: || |||
Db 7589 FALECVVTGTPELSAKFKDGRSADSKHITTFINKVASLKI PCAEMSDKGLYSFEVKN 7648
QY 92 -----FNNYVDGV----- 99
: || |||
Db 7649 SVGKSNCTSVHVSRIVPPSFIRKLDVNAILGASVVLECRVSGSAPISVGWFOGNEI 7708
QY 92 -----FNNYVDGV----- 99
: || |||
Db 7709 VSGPKQCSFSENVCNLNLSLEPSDTGIYTCVAANVAGSDECSAVLTVQEPPEQTPD 7768
QY 92 -----FNNYVDGV----- 99
: || |||
Db 7769 SVEVLPGHSLTFTSVIRGTPPKVFKWFGSRELYPGESCNI SLEDFTVELELFEVQPLES 7828
QY 92 -----FNNYVDGV----- 99
: || |||
Db 7829 GDYSLVTNDAGSACSTTHLFVKEPATFVKRLADF SVETGSPIVLEATYTGTPPISVSI 7888
QY 92 -----FNNYVDGV----- 99
: || |||
Db 7889 KDEYLISQSERCSITWTEKSTILESTIEDYAOYSCLIENRAGQDICEALVSVLEPPY 7948
QY 92 -----FNNYVDGV----- 99
: || |||
Db 7949 FIEPLEHVEAVIGEPATLQCKVDGTPPEIRISWYKEHTKLSAPAYKMQFKNNVASLVINK 8008
QY 92 -----FNNYVDGV----- 99
: || |||
Db 8009 VDSVDGEYCKADNSVGAVASSAVLVITKERKLPFPFFARKLKDVHETLGFVAFECRING 8068

QY 92 -----FNNYVDGV----- 99
: || |||
Db 8069 SEPLQVSWYKDGVLKDDANLOTFSFVNVTALQILQTDOSHIGQYNCSASNPLGTASSSA 8128
QY 100 -----FNNYVDGV----- 99
: || |||
Db 8129 KLILSEHEVPPFDLKPVSVDLALGESGTFKCHVTGTAPIKITWAKDNREIRPGGNYKT 8188
QY 100 -----FNNYVDGV----- 99
: || |||
Db 8189 LVENTATLTVLKVKGADAGQYTCYASNIAGKDCSAHLGVQBPFRFIKLEPSRIVKODE 8248
QY 100 -----FNNYVDGV----- 99
: || |||
Db 8249 FTRYECKIGGSPEIKVLWKDETEIQESSKFRMSFVDSVAVLEMHNL SVEDSGDYTCDAH 8308
QY 103 NA-----FNNYVDGV----- 104
: || |||
Db 8309 NAAGSASSSTSLKVKEPPIFRKKPHPIETLKCADVHLECELOQTPEFHVSWYKDKRELRS 8368
QY 105 -----FNNYVDGV----- 109
: || |||
Db 8369 GKYYKIMSENFILTSIHLNVDAADIGEYQCKATNDVSDTCVGSIALKAPRFRVKKLSDI 8428
QY 110 -----FNNYVDGV----- 109
: || |||
Db 8429 STTVGKEVOLQTTIEGABPISVWPFKDKGEIVRESNIIWISYSENIIATLQFSRVEPANAG 8488
QY 110 -----FNNYVDGV----- 109
: || |||
Db 8489 KYTCQIKNDAGMOECFATLSVLEPATIVEKPESIKVTTGDTCTLECTVAGTPELSTKWPK 8548
QY 110 -----FNNYVDGV----- 109
: || |||
Db 8549 DGKELTSNDKYSIPFNKVSGLKIINAPSDSGVYSFEVQNPVGKDSCTASLQVSDRTPV 8608
QY 110 -----FNNYVDGV----- 109
: || |||
Db 8609 PSFTRKLLKETNGLSGSSVMECKVYGSPPISVSWFHEGNEISSGRKKYQTTLTDTNCTALV 8668
QY 110 -----FNNYVDGV----- 111
: || |||
Db 8669 NMLESDSGDYTCIATNMAGSDECSAPLTVREPPSVQKPDPMVLTGTNTVFTTSIVKGT 8728
QY 112 -----FNNYVDGV----- 111
: || |||
Db 8729 PPFVSWFKGSSELVPGDRCNVSLSDSVALELFDVDT SQGEYTCIYSNEAGKASCTTH 8788
QY 112 -----FNNYVDGV----- 111
: || |||
Db 8789 LYIKAPAFVRLNDYSIEKGPLILEGTFGTGTPDISVTWKKNGINVTQSRQCNITTEK 8848
QY 112 -----FNNYVDGV----- 114
: || |||
Db 8849 SAILLEPSTVEDAGQYNYCIENASGKDCSAQILILEPPYFVKQLEPVKVSVDGDSASIQ 8908
QY 115 -----FNNYVDGV----- 119
: || |||
Db 8909 CQLAGTPEIGVSWYKGDTKLRPTTTTKMHFRNNVATLVFNQVDINDSGEYICKAENSUGE 8968
QY 120 -----FNNYVDGV----- 122
: || |||
Db 8969 VSASTFLTVOEQLPPPSRQLRDVQETVGLPVVFDCAISGSEPI SVSWYKDKPLKDSKP 9028
QY 123 -----FNNYVDGV----- 122
: || |||
Db 9029 NVQTSFLDNTATLNIFKTDRLAGQY SCTATNP IGSASSARLILTEGKNPPFFDIRLAP 9088
QY 123 -----FNNYVDGV----- 129
: || |||
Db 9089 VDVAVGESADFECHVTGTQPIKVSMAKDSREIRSGGKYQISYLENSAHLTVLKVDKGDSG 9148

QY	130	-----WLNKGE-----	135
Db	9149	QYTCYAVNEVGKDSCTAQLNIKERLIPPSFTKRLSETVEETEGNSFKLEGRVAGSQITV	9208
QY	136	-----YKCKVSN-----	142
Db	9209	AWYKNNIEIOPTNCEITFKNNTLVQVRKAGMNDAGLYTCKVSNDSAGSALCTSSIVIKE	9268
QY	143	-----	142
Db	9269	PKKPPVFDQHLTPVTYSEGEVQVLSCHVQSEPIRIQWLKAGREIKPSDRCSFSASGTA	9328
QY	143	-----	142
Db	9329	VLELRDVAKADSGDYVCKASNAVAGSDTTKSKVTIKDKPAVAPATKKAANDGRLFFVSEPQ	9388
QY	143	-----	142
Db	9389	SIRVVEKTATFIKVGGDIPNVKWTGKWROLNOGGRVFIHQKGDCAKLEIRDTTKTD	9448
QY	143	-----	142
Db	9449	SGLYRCVAFNEHGEIESNVNLQVDERKKQEKIEGDLRLMLKKTPIILKKGAGEEEDIME	9508
QY	143	-----	142
Db	9509	LLKNVDPKEYEKYARMYGITDFRGLLOAFELLKQSOBEETHRLEIEIERSERDEKEFEE	9568
QY	143	-----	142
Db	9569	LVSEIQRLSQTEPVLTIKDIENQTVLKNDVAFEDIKINYPEIKLSWYKGTLEKLEPSD	9628
QY	143	-----	142
Db	9629	KFEISIDGDRHLRLVKNCLQDOGNVLCVCGPHIASAKLTVIEPAWERHLQDVTLEKQGT	9688
QY	143	-----	142
Db	9689	CTMTCQFSVPNVKSEWFRNGRILKPOGRHKTEVEHKVHLTIADVRAEDQGYTCRYEDL	9748
QY	143	-----	142
Db	9749	ETSAELRIEAPIQTFKRIQNIIVVSEHQSATFECEVSFDDAIVTWYKGPTELTESOKYNF	9808
QY	143	-----	142
Db	9809	RNDGRCHYMTIHNWTPDDEGVSVIARLEPRGEARSTAEILYLTKEIKLEKPPDIPDSR	9868
QY	143	-----	142
Db	9869	VPITMPIRAVPPPEIPVAVAPPILPLPTPEEKPPPKRIEVTKKAVKDKAKKVAKPK	9928
QY	143	-----	142
Db	9929	EMTPREIIVKKPPPTTLIPAKAPEIDVSSKAEEVKIMTITRKKVEQKEAVEKKQA	9988
QY	143	-----	142
Db	9989	VHKEKRVIESPEPYDELEVEPYTEPFEOYYVEEDPEDEYEEKVKAKEVHEEWEEDFE	10048
QY	143	-----	142
Db	10049	EGOEYTEREGYDEGEENEWEAYQEREVIQVQKEVYESHERKVPAKVPEKKAPPPKVI	10108
QY	143	-----	142
Db	10109	KKPVIEKIEKTSRMEEEKVQVTKVPEVSKKIVPQKPSRTPVQEVIEVKVPAVHTKKMV	10168
QY	143	-----	142
Db	10169	ISEKMFASHTEEEVSVTVPEVQKEIVTEEKIHVAVSKRVEPPPKVPELPEKPAPEVA	10228
QY	143	-----	142

Db	10229	PVPIPKKVEPPAPKVPPEVPKKVPPEKKVPVPKKKEPAAPPKVPPEVKKVPPEEKIPVPV	10288
QY	143	-----	142
Db	10289	AKKKEAPPKVPPEVQKRVVTVEEKITIVTQREESPPPAVPEIPKPKKVPPEERKVPKKEEV	10348
QY	143	-----	142
Db	10349	PPPKVPALPKKVPPEEKVAVVPVPAKAPPPRAEVSKTIVVEEKRFVAEKLFSFVAPQR	10408
QY	143	-----	142
Db	10409	VEVTRHEVSAAEEWSYSEEEEGVSISVYREEREEREEBAEVEYVAMEEPEEYVVEEKLH	10468
QY	143	-----KALPAPI-----	149
Db	10469	IIISKRVEAEPAEYTEROEKKIVLKPPIAKIEEPPPAKVPPEAKKIVPEKKVPAPVPKKE	10528
QY	150	-----	149
Db	10529	KVPPKVPPEEPKVPPEKKVPPKVKMEEPPLPAKVTERHMQITQEEKVLVAVTKKEAPPK	10588
QY	150	-----	149
Db	10589	ARVPEEPKRAVPEEKVLKPKREEDPAKVTEFRKRVKKEKVSIEAPKREPOPIKEVT	10648
QY	150	-----	149
Db	10649	IMBEKERAYTLEBAVSVQREEEYEEYDYKEFEFEYETEEYDQYEEYEEYEEYEE	10708
QY	150	-----	149
Db	10709	HEEYITEPEKPIPVKVPPEEPVTKPAKPAKVLKKAPEEKVPVPIPKLKPDPKVPPE	10768
QY	150	-----	149
Db	10769	EPKKVFEKIRISITKREKQVTEPAKVPKPKRVAAEEKVPVPRKEVAPPVPRVPEVK	10828
QY	150	-----	149
Db	10829	ELEPEEVAFEVEVTVHVEEYLVVEEEYIHEEEEFITEEEVVPVPIVKVPEVPKVPPEE	10888
QY	150	-----	149
Db	10889	KKVPVPKKKEAPPKVPPEVPKKPEEKVPVLIIPKKEKPPPAKVPPEVKKVPPEEKVPVPV	10948
QY	150	-----	149
Db	10949	PKKVEAPPKVPPEVPKKVPPEKKVPVPAPKPKVEAPPKVPPEVKKLIPEKKKPTVPKVK	11008
QY	150	-----	149
Db	11009	EAPPKVPKKREVPVPVVALPQEEVLFEEIIVPEEVLPEEEVLPEEVLPEEVLPEEVL	11068
QY	150	-----	149
Db	11069	PEEEIIPPEEEVPPPEEYVPEEEFVPEEVLPEVKVPVPAPVPEIKKKVKTEKKVI	11128
QY	150	-----	149
Db	11129	PKKEAPPKVPPEVPKKVEKKRIILPKEEVLPEVEETPEEPEEPISEETPEEPPSIEEV	11188
QY	150	-----	149
Db	11189	EEVAPPRVPEVIKKAPEAPTVPKVEAPPKVSKKIPEEKVPVPVOKKEAPPKVPPEV	11248
QY	150	-----	149
Db	11249	PKKVPKKVLVPKKEAVPPAKGRTVLEEKYSVAFRQEVVVVKERLELEVVAEVEEPIEEE	11308
QY	150	-----	149

Db 11309 EFHEVEYFEGEFHEVEBEFIKLEOHRVBEHRVEKVRHVEIEVEAEVEVEFEKPKAPPK 11368
QY 150 ----- 149
Db 11369 GPEISEKIIIPPKPPKPPKVVPRKEPPAKVPEVPKKIIVEEKVRVPEPRVPPTKVPDVLPP 11428
QY 150 ----- 149
Db 11429 KEVVPKVPVPPAKKPEAPPKVPPEAPKEVVPKVPVPPPKKPEVPPTKVPPEVPPKAAV 11488
QY 150 ----- 149
Db 11489 PEKKVPEAIIPKPESPPEVPEAPKEVPEKKVPAAPPKKPEVTPVKVPEAPKEVVPK 11548
QY 150 ----- 149
Db 11549 VPPPPKPEVPPTKVPPEVKVAVPEKKVPEAIIPKPESPPEVPEPEPEVALEPPAEV 11608
QY 150 ----- 149
Db 11609 VEEPEAAPQVTPPKKVPPEKKAPAVAKKPELPPVKVPEVPKPEKKVPLVVPK 11668
QY 150 ----- 149
Db 11669 PEAPPKVPPEVPKPEKKVAVPKKPEVPPAKVPEVPKPVLEKPAVPPPERAESPPP 11728
QY 150 ----- 149
Db 11729 EYVEPEETAPEETAPESEKVPVNAEBEPEVPPPAVPEPEPKKIPEKKVPVKKPEAP 11788
QY 150 ----- EKTISK ----- 155
|||
Db 11789 PKPEPEKVEKPKLPPPPPPPPAPPKEDYKIFQLKAIKPKKVPKPOVPEKVELT 11848
QY 156 ----- 155
Db 11849 PLKVPGEKKVRKLLPERKPEPEEYVLKVRKPEEPEKPEKLEKVKKPAVPEPP 11908
QY 156 ----- 155
Db 11909 PPKPVEVEVPTTKRERKIPEPTKVPKPAIPALPAPPEPKKPEAEVKTIRPPVPEPP 11968
QY 156 ----- 155
Db 11969 TPAAAPVTPVVGKKAAPKAEAAKPKPIKVPKTPSPTEAERRKLKRGSGKEPP 12028
QY 156 ----- 155
Db 12029 DEAPFTYQLKAVPLKVPKVEIKDIILTESEFVGSSAIFECVSPSTAITMMKDGSNIRES 12088
QY 156 ----- 155
Db 12089 PKHRTADGDKRKLHIIDVQLSDAGEYTCVLRGNKEKTSTAKLVVEELPVRPVKLEEE 12148
QY 156 ---AKGP----- 160
||||
Db 12149 VTVWKGPLYLSELNKERDVVRKDKGKIVPEKPGRIVPGVIGLMRALTINDADDTAGT 12208
QY 161 ----- 160
Db 12209 YTVTVENANNLECCSKVKVVEVIRDLVKPIRDQHVKPKGTAFACDIATKDPNFKWFG 12268
QY 161 ----- 160
Db 12269 YDEIPAEPNDKTEILRDGNHLYLKIKNAMPEDIAEYAVEIEGKRYPAKLTIGEREVELLK 12328
QY 161 ----- 160
Db 12329 PIEDVTIYESASFDAEISEADIPQWKLKELLRPSPTCEIKAEGGKRLTLAKVKLD 12388
QY 161 ----- 160
Db 12389 QAGEVLYQALNAITTAILTVEIELDFAVPLKDVTPERRQARFECVLTRANVTWSKGP 12448

QY 161 ----- 160
Db 12449 DIIKSSDKFDIIADGKKHILVINDSQFDDGEGVYTAEBEGKTSARLFTVIGIRLKFNPSPLE 12508
QY 161 ----- 160
Db 12509 DQTVKEGETATFVCELSHEKMHVWPKNDAKLHTSHTVLJSSGKTHKLEMEVTLDDLS 12568
QY 161 ----- 160
Db 12569 QIKAQVKELSTAQLKVLADPYFTVKLHDKTAVEKDEITLKCEVSKDVPVKWFKDGBEI 12628
QY 161 ----- 160
Db 12629 VPSPKYSIKADGLRILKIKKADLKDKGVEYVCDGTDKTKANVTVEARLIKVEKPLYGVE 12688
QY 161 ----- 160
Db 12689 VVGETAHEIEIELSEBPDVHGQWKLKQPLTASPDCEIIEDGKKHILILHNCQLGMTGEYS 12748
QY 161 ----- REPQ ----- 164
|||
Db 12749 FOANAKSAANLKVRELPLIFITPLSDVKVFEKDEAKFECEVSREPKTPRWLKGTOEITG 12808
QY 165 ----- 164
Db 12809 DDRFELIKDGTKHSVVIKSAAFEDAKYMFEAEDKHTSGKLLIEGIRLKFPLKDVATAK 12868
QY 165 ----- 164
Db 12869 EKESAVFTVELSHDNIRVKWFKNDQRLHTRSVSMQDEGKTHSITFKDLSIODTSIRVE 12928
QY 165 ----- 164
Db 12929 AMGMSSEAKLTVLEGDPYFTGKLQDYTGVEKDEVILQCEISKADAPVKWFKDGKEIKPSK 12988
QY 165 ----- 164
Db 12989 NAVIKADGKKRMLILKKALKSDIGQYTCDCGTDKSGKLDIEDREIKLVRPLHSVEVMT 13048
QY 165 ----- 164
Db 13049 ETARFETEISDDIHANWKLKGEALLQTPDCEIKBEGKIHSVLHNCRLDQGTGVDFQAA 13108
QY 165 ----- 164
Db 13109 NVKSSAHLRVKPRVTIGLLRPLKDVTVTAGETATFDCELSYEDI PVWYLGKKLEPSDKV 13168
QY 165 ----- 164
Db 13169 VPRSEKVTTLTRDVKLEDAGEVOLTAQDFKTHANLFVKPEPPVFTKPLEDQTVEEGAT 13228
QY 165 ----- 164
Db 13229 AVLECEVSRENKAKVWFKNGTEILKSKKYEIVADGRVRKLVIHDTCTEDIKTYTCDAKDF 13288
QY 165 ----- 164
Db 13289 KTSCLNVVPPHVEFLRPLTLQVREKEMARFECELSRENKAKVWFKDGAIEKKKKYDI 13348
QY 165 ----- 164
Db 13349 ISKGAVRILVINKLLDDEAEYSCEVRTARTSGMLTVLEEEAVFTKNLANIEVSETDTIK 13408
QY 165 ----- 164
Db 13409 LVCEVSKPGAIEVIWYKGDEEIIETGRYEILTEGRKRILVIONAHLEDAGNVCRLPSSRT 13468
QY 165 ----- 164
Db 13469 DCKVKVHELAAEFISKPONLEILEGEKAEFVCSISKSEFPQWKRDDKTLESGDKYDVITA 13528

QY 165 ----- 164
Db 13529 DGKRVLVKDATLQDMGTIVVMVGARAAAHLTVEKLRIVVPLKDRVKEQEVFNC 13588
QY 165 -----VYTL----- 168
Db 13589 EVNTEGAKAKFRNEAIFDSSKYIILQKDLVYTLRDAHLDDQANYVSLNHRGENV 13648
QY 169 ----- 168
Db 13649 KSAANLIVEEDLRIVEPLKDIETMEKSVTFWCKVNRNLNVTILKWTNGEVEVFFDNRVSY 13708
QY 169 ----- 168
Db 13709 RVDKYRKMILTICKGFPDEGEYIVTAGQDKSVAELLIIEAPTEFVEHLEDQTVTEFDDAV 13768
QY 169 ----- 168
Db 13769 FSQLSREKANVKWYRNGREIKEGKKYKFEKDGSIHRLIICKDRLDCEYACGVEDRKS 13828
QY 169 ----- 168
Db 13829 RARLFVEEIPVEIIRPPQDILEAPGADVFLAELNKKDVEQVLRNMMVVVQDKHOMMS 13888
QY 169 ----- 168
Db 13889 EGIHRLOICDIKPRDQGEYRFIAKDEARAKLELAAPKIKTADQDLVVDVGKPLTMV 13948
QY 169 ----- 168
Db 13949 PYDAYKAEAEWFKENEPLSTKTIDTTAEQTSFRILEAKKGDKGRYKIVLQNKHKAEGF 14008
QY 169 ----- 168
Db 14009 INLKVIDVPGVRNLEVTFETFDGEVSLAWEELTDGSKLIIGYVVERRDIKRTWVLATD 14068
QY 169 ----- 168
Db 14069 RAESCEFTVTGLQKGVEYLFVRSARNRVGTGEPVETDNPVEARSKYDVPGLNVTID 14128
QY 169 ----- 168
Db 14129 VNRFGVSLTWEPPEYDGGAEITNYVIELRDKTSIRWDATMTVRAEDLSATVTDVVEGEY 14188
QY 169 ----- 168
Db 14189 SFRVRAQNRIGVGKPSAATPFVKVADPIERPSPVNLTSDDQTQSSVOLKWEPPKLDGGS 14248
QY 169 ----- 168
Db 14249 PILGYIIERCEEKDNWIRCNMMLVPELITYKVTGLEKNKLYRVSANENKAGVSDPSEIL 14308
QY 169 ----- 168
Db 14309 GPLTADDAFVEPTMDLSAFKDGLEIVPNPITILVPSTGYPRPTATWCFGKVLGTGDRV 14368
QY 169 ----- 168
Db 14369 KMTLSAYALVISPERSDKGIYTLKLENRVKTIISGEIDVNVIARPSAPKELKFGDITK 14428
QY 169 ----- 168
Db 14429 DSVHLTWEPDDGGSLTGYYVEKREVSRTWTVMDFVTDLEFTVVDLVQGEYLFKV 14488
QY 169 ----- 168
Db 14489 CARNKCGPGEPAYVDEPVNMSTPATVPDPPENVKWRDRTANSIFLTWDPKNDGGSRIKG 14548
QY 169 ----- 168
Db 14549 YIVERCPRGSKWVACGEPVAETKMEVTGLEBGKWYAYRVKALNRQAGSKPSRPTBIEQA 14608
QY 169 ----- 168

Db 14609 VDTQEAPEIFLDVYKLLAGLTVKAGTKIELPATVTGRPEPKITWTKADMILKODKRITEN 14668
QY 169 ----- 168
Db 14669 VPKKSTVTIVDSKRSDTGTYYIEAVNVCGRATAVVENVLDKPGPPAAFDITDVTNESCL 14728
QY 169 ----- 168
Db 14729 LTNWPRDDGGSKITNVVVERRATDSEVWHKLSSTVKDTNFKATKLIIPNKEYIFRVAAN 14788
QY 169 ----- 168
Db 14789 MYGVGEVQASPIAKYQFDPGPPTRLEPDSITKDAVTLTWCEPDGDSPTIGYWVER 14848
QY 169 ----- 168
Db 14849 LDPTDKWVRNCNMPVKDITYRVKGLTNKKYRFRVLAENLAGPGPKSPKSTELILKDPI 14908
QY 169 ----- 168
Db 14909 DDPWPFKPTVKDVGKTSVRLNWTKPEHDGGAKIESVIEMLKTGTDWVRVAEGVPTTQ 14968
QY 169 ----- 168
Db 14969 HLLPGLMEGOEYSFRVRAVNAKAGESBPSPVLCREKLYPPSPRWLEVINITKNTAD 15028
QY 169 ----- 168
Db 15029 LKWTVEKGGSPITNVIIEKRDVRRKGWQTVDTTKTKCTVTPLTEGSLYVFRVAAN 15088
QY 169 ----- 168
Db 15089 AIGQSDYTEIEDSVLAKDTFTTTPGPPVALAVDVTKRHVDLKWEPPKNDGGRPIQRYVIE 15148
QY 169 ----- 168
Db 15149 KKERLCTRWYKAGTAGPCNFRVTDVIEGTEVQFQVRAENEAGVGHSPSEFTEILSIEDP 15208
QY 169 ----- 168
Db 15209 TSPSPPLDLHVTDAGRKKHIAIAWKPEKNGGSPITIGYHVEMCPVGTCKMWRVNSRPIDK 15268
QY 169 ----- 168
Db 15269 LKFKVEEGVVPDKKEYVLRVRAVNAIGVSEPISENVAIVAKDPCKPTIDLETHDIIIVIEG 15328
QY 169 ----- 168
Db 15329 EKLSIPVPFRAPVPVTVSWHKDGKEVKASDRLTMKNDHISAHLEVPKSVRADAGIYITL 15388
QY 169 ----- 168
Db 15389 ENKLSATASINVKVIGLPGPKDIAKASDITKSSCKLTWEPPEFDGTPILHYVLERREA 15448
QY 169 ----- 168
Db 15449 GRRTYIPVMSGENKLSWTVKDLIPNGEYFFRVRKAVNKVGGGEYIELKNPVIADQPKQPPD 15508
QY 169 ----- 168
Db 15509 PPVDVEVHNPTAEAMTITWKPPLYDGGSKIMGYIIEKIAKEERWKRNEHLVPLTYTA 15568
QY 169 ----- 168
Db 15569 KGLECKEYQFVRAENAGISEPSRATPPTKAVDPIDAPKVILRTSLEVKRGDEIALDA 15628
QY 169 ----- 168
Db 15629 SISGSPYTTITWIKDENVIIVPEETIKKRAAPLVRRRKCEVQEEFPVPLTQRLSIDNSKK 15688
QY 169 ----- 168

Db 15689 GESQVRDRLRPDHLGYMIKVENDHGIAKAPCTVSLDTPGPPINFVEDIRKTSVLCK 15748
QY 169 ----- 168
Db 15749 WEPPLDDGGSEIINTLEKKDKTKPDSEWIVVTSLRHCKYSVTKLIEGKEYLFRVAEN 15808
QY 169 ----- 168
Db 15809 RFGPGPCYSKPLVAKDPFGPDADPKPIVEDVTSNSMLVKWNEPKDNGSPILGYWLEKR 15868
QY 169 ----- 168
Db 15869 EYNSTHWSRVNKSLLNALKANDVGLLEGLTYVFRVCAENAAGPKFSPSPDKTAHDPTS 15928
QY 169 ----- 168
Db 15929 PPGPPIPRVTDSSTTIELEWEPAPNGGGEIVGFVDKLVGTNWSRCTEKMVKVROY 15988
QY 169 ----- 168
Db 15989 TVKEIREGADYKLRSVAVNAAGEPPGETQPVTVAEPOEPPAVELDVSVKGGIIMAGKT 16048
QY 169 ----- 168
Db 16049 LRIPAVVTGRPVPTKVTWKEGELDKDRVVDNVGTKSELIIKDALKRDHGRVITATNS 16108
QY 169 ----- 168
Db 16109 CGSKFAARVEFDPGVPVLDLKPVTNPKMCLLNWSDPDDGGSEITGFIIERDKAMH 16168
QY 169 ----- 168
Db 16169 TWRQPIETERSKCDITLLEGOEYKFRVIANKFKGPGPVEIGPILAVDPLGPPTSPEL 16228
QY 169 ----- 168
Db 16229 TYTERTKSTITLDWKEPRNSGSPIOGYIIEKRHRDKPDFERNKRLCPTTSFLVENLDE 16288
QY 169 ----- 168
Db 16289 HQMYEFRKAVNEIGESPSPLNVVIOQDEVPPTIKLRSVRGDTIKVAGEPVHIAD 16348
QY 169 ----- 168
Db 16349 VTGLPMPKIEWSNETVIEKPTDALQITKEYSRSEAKTELSIPKAVREDKGYTVVTASN 16408
QY 169 ----- 168
Db 16409 RLGSVFRNVHVEYDRPSPPRMLAVTDIKAESCYLTDAPLDNGGSEITHYVIDRKDASR 16468
QY 169 ----- 168
Db 16469 KKAEEVTNTAVEKRYGIWKLIPNGQYEFVRVAVNKKYGISDECKSDKVVIOQPYRLPGP 16528
QY 169 ----- 168
Db 16529 PGKPKVLARTKGLMSVSWTPPLDNGSGPITGYWLEKREBEGSPYWSRVSRAPIITKVGLKGV 16588
QY 169 ----- 168
Db 16589 EFNVPRLLEGVYQFRAMAINAAGIGPPSEPSDPEVAGDPIFPGPPSPCEVKDKTSSI 16648
QY 169 ----- 173
Db 16649 SLGWKPPAKDGGSPIKGYIVEMQEBGTTDWKRVNEPDKLITTCCEVVPNLKELRYFRV 16708
QY 174 ----- 173
Db 16709 KAVNEAGESEPSDITGTGIEPATDIOBEPEVFDIGAQDCLVCKAGSOIRIPAVIKGRPTPK 16768
QY 174 ----- 173
Db 16769 SSWEPDGAKKAMKGVHDIPEDAQLETAENSSVIIPECKRSHTGKYSITAKNAGOKT 16828

QY 174 ----- 173
Db 16829 ANCRVKVMDVPGPPKDLKVSDITRGSCLSWKMPDDDDGGDRIKGYVIEKRTIDGKAWTKV 16888
QY 174 ----- 173
Db 16889 NPDCGSTTFVVPDLLSEQQYFFRVRRAENRFGIGPPVETIQTARTDPIYPPDPPIKLGIG 16948
QY 174 ----- 173
Db 16949 LITKNTVHLSWKPPKNDGGSPVTHYIVECLAMDPTGKKEANRQCNRDVEELQFTVEDL 17008
QY 174 ----- 173
Db 17009 VEGGEYFRVKAUNAAGVSKPSATVGPCDCQRPDPSPSIDLKEFMEVEEGTNNVIVAKIK 17068
QY 174 ----- 173
Db 17069 GVPFPTLTWFKAPKKPDNKEPVLVDYTHVNLVVDCTFLVIPSRRSDTGLVITAVNN 17128
QY 174 ----- 173
Db 17129 LGTASKEMLNVLGRPPPVGPIKFSVSADQMTLSWPPKDDGSGKITNYVIEKREANR 17188
QY 174 ----- 173
Db 17189 KTWVHVSSEPKRECTYTIKLLLEGHEVPRIMAQNKYIGIGEPDLSBPETARNLFSVPGAPD 17248
QY 174 ----- 173
Db 17249 KPTVSSVTRNSMTVNWBEPEYDGGSPVTGYWLEMDKTTSKRWKVRNRPDIKAMTLGVSYK 17308
QY 174 ----- 173
Db 17309 VTGLIEGSDYQFRVYAINAAGVGPASLPSPATARDPIAPGPPPKVTDWTKSSADLEW 17368
QY 174 ----- 173
Db 17369 SPPLKGGSKVTGYIVEYKEEGKEWEKDKVEKGTKLVTGLKEGAFYKFRVSAVNIA 17428
QY 174 ----- 173
Db 17429 GIGEPGEVTDVTEMKDLVSPDLQLDASVRDRIVVHAGGVIIRIIAYVSGKPPPTVTWNNN 17488
QY 174 ----- 173
Db 17489 ERTLPOEATIIETAISSSWIKNCORSHQGVYSLAKNEAGERKKTIIIVDVLDPGPVCT 17548
QY 174 ----- 173
Db 17549 PFLAHLNLTNESCCLTWFSPEDDGGSPITNYVIEKRESDRRAWTPVTYIVTRQATVQGLJ 17608
QY 174 ----- 183
Db 17609 QGKAYFFRIIAENSIGMGPFVETSEALVIREPITVPERPEDLEVKEVTKNTVTLTWNPCK 17668
QY 184 ----- 183
Db 17669 YDGGSEIINYVLESRLIGTEKFKVTNDNLLSRKYTVKLGEGDTYFYRVSAVINVGOK 17728
QY 184 ----- 183
Db 17729 PSFCTKPTCKDELAPPTLHLDLFRDKLIRVGEAFALTGRYSGKPKPKVSWFKDEADVLE 17788
QY 184 ----- 189
Db 17789 DDRTHIKTTPATLALEKIKAKRSDSGKYCVVVENSTGSRKGCQCVNVVDRPGPPVPSF 17848
QY 190 ----- 189
Db 17849 DEVTKDYMVISWKPPLDDGGSKITNYIIEKKEVKDVMVMTSASAKTCKVSKLLEGKD 17908

QY 190 ----- 189
Db 17909 YIFRIHAENLYGISDPLYSDSMKAKDRFRVPDAPDQPIVTEVTKDSALVTWNKPHDGGKP 17968
QY 190 ----- 189
Db 17969 ITNYILEKRETMKSRWARVTKDPIHPYTKFRVPDLLGCGQYEFVRSANEIGIGDPSPPS 18028
QY 190 ----- 189
Db 18029 KPVFAKDPIAKSPVPNPEALDITCNSVDLTWOPPRHDGSKILGIYIVEYKVGDEWRR 18088
QY 190 ----- 189
Db 18089 ANHTPESCPEYKYVTKGLRDGQTYKFRVLAVNAAGESDPAHPVPEVLVKDRLEPPELILD 18148
QY 190 ----- 189
Db 18149 ANMAREQHIKVGDTLRLSALIKGVPPPKVTWKKEDRDAPTAKRIDVTPVGSKLEIRNAAH 18208
QY 190 ----- 189
Db 18209 EDGGIYSLTVENPAGSKTVSKVLVLDKPGPRDLEVSEIRKDCYLTKWEPLDDGGSVI 18268
QY 190 ----- 189
Db 18269 TNYVVERRDVASQWSPLSATSKKKSHFAKHLNEGNOYLFERVAENQYGRGPFVETPKPI 18328
QY 190 ----- 189
Db 18329 KALDPLHPPGPKDLHHVDKTEVSLVWNKPRDRDGGSPITGYLVEYQESTQDWIKFKT 18388
QY 190 ----- 189
Db 18389 VTNLECVVTGLQOKTYFRVRAENIVGLGLPDTTIPIECOEKLVPSPVELDKLIEGLV 18448
QY 190 ----- 189
Db 18449 VKAGTTVRFPAILIRGVVPPTAKWTTDGGSEIKTDEHYTVETDNFSSVLTIKCLRRDTGEY 18508
QY 190 ----- 189
Db 18509 QITVSNAAKSTVAVHLTVLDVPGPTGPIINILDVTEHMTISWQPPKDDGGSPVINYIV 18568
QY 190 ----- 189
Db 18569 EKQDTRKDTWGVSSGSKTKLIPHLOKGEYVFRVRAENKICVGPPLDSTPTVAKHKF 18628
QY 190 ----- 189
Db 18629 SPPSPGKPVVTDITENAAATVSWTLPKSDGGSPITGYMERREVTGKWRVKNKPTIADLK 18688
QY 190 ----- 189
Db 18689 FVRTGLYEGNTYEFVFAENLAGLSKPSDDPIKACRPKPPGPPINPKLDKKSRETAD 18748
QY 190 ----- 189
Db 18749 LVWTKPLSDGSPILGVVVEQCPGTAQWNINKDELIROCAFVPGGLIEGNEYFRFIKA 18808
QY 190 ----- 189
Db 18809 ANIVGEPELAEVIAKDILHPPEVELDVTCRDVITVRVGTIRILARVKRPEPDIT 18868
QY 190 ----- 189
Db 18869 WTKGKVLVREKRVLDLIQDLPRVELQIKEAVRADHGKYIISAKNSSGHAOGSAIVNVLDL 18928
QY 190 ----- 189
Db 18929 PGPCQNLKVTNVTKENCTISWENPLDNGSGSEITNFIVEYRKPQKQHSIVASDVTKKLK 18988
QY 190 ----- 189

Db 18989 ANLLANNEYFRVCAENKVGVPITETKTPIALINPIDRGPENLHIADKGKTFVYLKW 19048
QY 190 ----- 189
Db 19049 RRPDYDGGSPNLVSHVERRLKGDDWERVHKSGIKETHYMWDRCVENQIYEFVQTKNEG 19108
QY 190 ----- 189
Db 19109 GESDWVKEEVVVKEDLQKPVLDLKLGSVLTVKAGDTIRLEAGVRGKPFPEVAMTKOKDA 19168
QY 190 ----- 189
Db 19169 TDLTRSPRVKIDTRADSKFSLTAKRSDGKYVVTATINTAGSFVAYATVNVLDKPGVPR 19228
QY 190 ----- 189
Db 19229 NLKIVDSSDRCTVCWDDPPEDDGGCEIQNYILEKCEIKRMVWSTYSATVLTPTGTVTVRLI 19288
QY 190 ----- 189
Db 19289 EGNEYIFRVRAENKIGTPPTESKPIAKTKYDKGRDPPEVTKVSKEMTVVWNPPEY 19348
QY 190 ----- 189
Db 19349 DGGKSITGYELEKKEKHSRWPVYNKSAIPERRMKVONLLPDHEYQFRVKAENEIGEP 19408
QY 190 ----- 189
Db 19409 SLSPRPVAKDPIEBPPGPPNFRVVDTTKHSITLGWGKPVYDGGAPIIGYVVMRPKIAD 19468
QY 190 ----- 189
Db 19469 ASPDEGKRCNAAQLYRKEFTVTLSDENQEYEFVCAQNOVGIGRPAELKEAIKPKEL 19528
QY 190 ----- 189
Db 19529 EPPEIDLDSMRKLVIVRAGCPIRLFAIVRGRPAKVTWRKVGIDNVVRKGQVDLDTMA 19588
QY 190 ----- 189
Db 19589 FLVIPNSTRDDSGKYSLTLVNPAGEKAVFVNVVRVLDTPGVPVSDLKVSQVTKTSCHVSWAP 19648
QY 190 ----- 189
Db 19649 PENDGGQVTHYIVEKREADRKTWSTVTPYVKTSFHVTNLVPGNEYFRVTAVNEYGPG 19708
QY 190 ----- 189
Db 19709 VPTDVPKPVLASDPLSEPPRPKLEVTMTKNSATLAWLPLRDGGAKIDGYITSYREE 19768
QY 190 ----- 189
Db 19769 QPADRWTEYSVVKDLSLVVTGLKEGKKYKFRVAARNAVGSVLPREAEGVYEAKEQLLPPK 19828
QY 190 ----- 189
Db 19829 ILMPEQITIKAGKKLRIEAHVYGRPHPTCKWKKEDEVVTSLSHLAVHKADSSSILIIKDV 19888
QY 190 ----- 189
Db 19889 TRKDSGYISLTAENSSGTDQKIKVVVMDAPGPPPPDISDIDADACSLSWHIPLEDGG 19948
QY 190 ----- 189
Db 19949 SNITNYIVEKCDVSRGDWVTALASVTKTSRCKLIPQOEYIFRVRAENRFGISEPLTSP 20008
QY 190 ----- 189
Db 20009 KWAQFPFGVPSEPKNARVTKVNKDCIFVADWRPDSGSGSPIGYLIERKERNILLSWKA 20068
QY 190 ----- 189

Db 20069 NDTLVRSTEYPCAGLVEGLEYSRIYALNKNKAGSSPPSKPTEYVTARMPVDDPPGKPEVIDV 20128
QY 190 ----- 189
Db 20129 TKSTVSLIWARPKHDGSKIIGYFVEACKLPGDKWVRCNTAPHQIPOEEYTATGLEEKAQ 20188
QY 190 ----- 189
Db 20189 YOFRAIARTAVNISPPSESDPVTILAENVPRIDLVSAMKSLITVKAGTNVCLDATVFG 20248
QY 190 ----- 189
Db 20249 KPMPVSWKKDGTLLKPAEIGIKWAMORNLTLELFSVNRKDSGDYTTITAENSSGSKSATI 20308
QY 190 ----- 189
Db 20309 KLVLDKPGPASVKINKMYSDRAMLSWEPLEDGSGSEITNYIVDKRETSRPNWAQVSAT 20368
QY 190 ----- 189
Db 20369 VPITSCSVEKLIEGHEYQFRICAENKYGDPVFTEPAIAKPNYPDPGRCDDPPVISNITK 20428
QY 190 ----- 189
Db 20429 DHMTVSWKPPADGGSPITGYLLEKRETOAVNWTKNRKPPIERTLKATGLOEGTEYEPR 20488
QY 190 ----- 189
Db 20489 VTAINKAGPGKSDAKAAYARDQYPPAPPAPPKYDTRSSVLSWCKPAYDGGSPIT 20548
QY 190 ----- 189
Db 20549 GYLVEYKRADSDNWVRNLPQNLOKTRFEVTGLMEDTOYQFRVYAVNKIGYSDPSDVPDK 20608
QY 190 -YPSDIIV----- 196
Db 20609 HYPKDLIPPEGELDADLRKTLILRAGVTMRLYVPVKGRPPPKITWSKPNVNLDRIGLD 20668
QY 197 ----- 196
Db 20669 IKSTDFTLRCENVNKYDAGKYILTLENSCGKKEYTIVVKYLDTPGPPVNVTVKEISKD 20728
QY 197 ----- 196
Db 20729 SAYVTWEPPIIDGSPFIINYVQKRAERKSWSTVTTECKTSFRVANLEEGSYFFRVF 20788
QY 197 ----- 196
Db 20789 AENEYIGDPGETROAVKASQTPGPPVVDLKVRSVKSSCSIGWKKPKHSDGGSRIIGYVVD 20848
QY 197 ----- 196
Db 20849 FLTEENKQVRMKSLSLQYSAKDLTEGKEYTFRVSAENENGEGTPEITVVARDDDVAPD 20908
QY 197 ----- 196
Db 20909 LDLKGLPDLCLAKENSFRKIPKIGKPAPSVSNKKGEDPLATDTRVSESSAVNTTLI 20968
QY 197 ----- 196
Db 20969 VYDQKSDAGYTTITLKNVAGTKEGTISIKVVVKPGIPTGPIKFDVTAEMTLKWAPPK 21028
QY 197 ----- 196
Db 21029 DGGSEITNYILEKRDVNNKHWTCASAVQKTTFRVTRLHEGMEYTFRVSANENKYGVEG 21088
QY 197 ----- 196
Db 21089 LKSEPIVARHPDVPDAPPPNIVDVHRDSVSLTWTDPKKTGSPITGVHLEFKERNILL 21148
QY 197 ----- 196
Db 21149 WKRANKTPIRMDFKVTGLTEGLEYEFRVMAINLAGVKPSLPSEPVVALDPIDPPGKE 21208

QY 197 ----- 196
Db 21209 VINITRNSVTLIWTEPKYDGGHKLTCYIVEKRDLPKSNMKNHVNVPCEAFTVTDLVEG 21268
QY 197 ----- 196
Db 21269 GKYEFRIAKNTAGAISAPSESTETIICKDEYEAPTIVLDPTIKDGLTIKAGDTIVLNAI 21328
QY 197 ----- 196
Db 21329 SILGRPLPKSSWSKAGKDIRPSDITQITSTPTSSMLTIKYATRCKDAGEYTTITATNPFGTK 21388
QY 197 ----- 196
Db 21389 VEHVKVTLVDPFGPGPVEISNVSAAKATLTWTPPLEDGGSPIKSYILEKRETSRLWTV 21448
QY 197 ----- 196
Db 21449 VSEDIQSCRHVATKLIQONEYIFRVSANVHYGKGEVPQSEPVKMVDRFGPPGPEKPEVS 21508
QY 197 ----- 196
Db 21509 NVTKNATVSWKRPVDDGGSEITGYHVERREKSLRWVRAIKTPVSDLRCKVTGLQEGST 21568
QY 197 ----- 196
Db 21569 YEFVSAENRAGIPPPSEASDSVLMKDAAYPPGPPSNPHVTDTTKKSASLAWGKPHYDGG 21628
QY 197 ----- 196
Db 21629 LEITGYVVEHOKVGEAWIKDITGTALRITQFVVPLQTKERYNFRISAINDAGVGEPAV 21688
QY 197 ----- 196
Db 21689 IPDVEIVEREMAPDFELDAELRRTLWVRAGLSIRFVPIKGRPAPEVTWTKONILKNRA 21748
QY 197 ----- 196
Db 21749 NIENTESFTLLIPECNRYDTGKFVMTIENPAGKKSGFVNVRVLDTPGPNLNRPTDITK 21808
QY 197 ----- 196
Db 21809 DSVTLHWDLPLIDGSRITNYIVEKREATRKSYSATTCHKCTYKVTVGLSEGEYFFRV 21868
QY 197 ----- 196
Db 21869 MAENEYIGEPTETTEPEVKASEAPSPDLSNIMDITKSTVSLAWPKPKHGGSKITGYVI 21928
QY 197 ----- 196
Db 21929 EAQRKGDQWTHITTVKGLECVVRNLTEGEYTFQVMVNSAGRSAPRESRPVIVKEQTM 21988
QY 197 ----- 196
Db 21989 LPELDLRIYQKLVIAKAGDNKIKVEIPVLRPKPTVTWKGDQILKQTVNFETTATST 22048
QY 197 ----- 196
Db 22049 ILNINECVRSDSGPPLFARNIVGEVGDVITQVHDIPGPTGPIKFDVSDFTFSWD 22108
QY 197 ----- 196
Db 22109 PPENDGGVPISNVVVMRQTDSTTWELATTVIRITYKATRLTTGLEYOFRVKAQNRVCV 22168
QY 197 ----- 196
Db 22169 GPGITSACIVANYPKVPVPGPTQVAVTKDSMTISWHEPLSDGSGPILGYHVERKERN 22228
QY 197 ----- 196
Db 22229 GILWOTVSKALVPGNIFKSSGLTDGIAVEFRVIAENMACKSKPSKPSEPMALDIPDPPG 22288

QY	197	-----	196
Db	22289	KVPLNITRHTVILKWKPEYTGFKITSYIVKRDLPNGRWLKNFNSILENEFTVSGL	22348
QY	197	-----	196
Db	22349	TEDAAVEFRVIAKNAAGISPPSEPSDAITCRDDVEAPKIKVDVKFDTVILKAGEAFL	22408
QY	197	-----EWESNGO-----	203
Db	22409	EADVSGRPPPTMEWSKDGELEGTAKLEIKIADFSTNLVKNKSTRDSDGAYTITATNPGG	22468
QY	204	-----	203
Db	22469	FAKHIFNVKLDLDRPGPEGLAVTEVTSEKCVLSWFPPLDDGGAKIDHYIVQKRETSRLA	22528
QY	204	-----	203
Db	22529	WTVASEVOVTKLKVTKLKKGNEYIFRMAVNKYGVGELESEFVLAVNPYGPDPKPN	22588
QY	204	-----	203
Db	22589	EVTIITKDSMVWCWGHDPDSDGSGEIIINYIVERRDKAGORWIKCNKKTLDLRYKVSGLTE	22648
QY	204	-----	203
Db	22649	GHEVEPRIMAENAGISAPSTSPFYKACDTVFKPGPPGNPRVLDTSRSSISIAWNKPIY	22708
QY	204	-----	203
Db	22709	DGSEITGYMVEIALPEDEWQIVTPPAGLKATSYTITGLTENQEKIRIYAMNSEGLGE	22768
QY	204	-----	203
Db	22769	PALVPGTPKAEDRMLPPEIELDADLRKVVTIRACCTLRLEVPKGRPAPEVKWARDGES	22828
QY	204	-----	203
Db	22829	LDKASIBESTSYTLLIVGNVNRFDGKYILTVENSNGSKSAFVNRVLDTPGPPQDLKVK	22888
QY	204	-----	203
Db	22889	EVTKTSVTLTWDPPPLDGGSKIRKINIVEKRESTRKAYSTVATNCHKTSWKVDQLQEGCSY	22948
QY	204	-----	203
Db	22949	YFRVLAENEYIGLPAETAESVKASERPLPPGKITLMDVTRNSVSLSWKPEHDGGSRL	23008
QY	204	-----	203
Db	23009	GIIVEMQTKGSKWATCATVKVTEATITGLIQGEYSFRVSAQNEKIGSDPRQLSVPVIA	23068
QY	204	-----	203
Db	23069	KDLVPPAFKLLFNFTVLAGEDLKVDPVFIGRPTPAVTWHKONVPLKQTRVNAESTEN	23128
QY	204	-----	203
Db	23129	NSLLTIKDACREDVGHVVKLTNSAGEAIEIETLVIVLDKPGPTGPVKMDEVTAADTSL	23188
QY	204	-----	203
Db	23189	WGPPKYDGGSSINNYIVEKRDSTTTWQIVSATVARTTIKACRLKTGCEYQFRIAENRY	23248
QY	204	-----	203
Db	23249	GKSTYLNSEPTVAQYKPKVPGPPTPVVTLSSRDSMEVQWNEPISDGGSRVIGYHLERKE	23308
QY	204	-----	203
Db	23309	RNSILWVKLNKTPIPQTKFTTGLEGEVEYFRVSAENIVGICKPSKVSECYVARDPCDP	23368
QY	204	-----	203

Db	23369	PGRPEALIVTRNSVTLQWKKPTYDGGSKITGYIVEKKELPEGRMKASFNTIIDTHEFTV	23428
QY	204	-----	203
Db	23429	GLVEDHRYEPRVIAARNAAGVSEFSESTGAITARDEVDPPRISMDPKYKDTIVVHAGESF	23488
QY	204	-----	203
Db	23489	KVDADIYGRPIPTIOWIKGDOELSNARLEIKSTDFATSLSDKAVDVDSNGVILKAKNV	23548
QY	204	-----	203
Db	23549	AGERSVTNVKVLDRPGPEGPVVISGVTAEKCTLAWKPLQDGGSDIINYIVERRETSR	23608
QY	204	-----	203
Db	23609	LWTVVDANVTLSCKVKTKLLEGNEYTRIMAVNKYGVGEPLSEFVPAKNPFVVPDAPK	23668
QY	204	-----	203
Db	23669	APEVTTVTKDSMIVVWERPASDGGSEILGYVLEKROKEGIRWTRCHKRLIGELRLRVGL	23728
QY	204	-----	203
Db	23729	IEHNDYFRVSAENAGLSEPPSPSAVQKACDPIYKPGPNPKVIDITRSSVFLSWSKP	23788
QY	204	-----	203
Db	23789	IYDGCIEIOGYIVEKCDVSYGEMTCTPTGINKNTNIEVEKLEKHEYNFRICAINKAGV	23848
QY	204	-----	203
Db	23849	GEHADVPPIIVEEKELEAPDIDLLELRKIINIRAGSLBLFVPIKGRPTPEVKWKVDG	23908
QY	204	-----	203
Db	23909	EIRDAALIDVTSFTSLVDNVNRDYGKYTLTLENSSGTKSAFVTVRVLDTSPVPVNLK	23968
QY	204	-----	203
Db	23969	VTEITKDSVSIWEPPPLDGGSKIKNYIVEKREATRKSAAVVTNCHKNSWKIDQLOEGC	24028
QY	204	-----	203
Db	24029	SYFVRVTAENEYIGLPAQTADPIKVAEVPQPPCKITVDDVTNRNSVSLSWTKPEHDGSK	24088
QY	204	-----	203
Db	24089	IIQYIVEMOAKHSEKSEKARVKSLOAVITNLTOGEEYLFVRVAVNEKGRSDPRSLAVPI	24148
QY	204	-----	203
Db	24149	VAKDLVTEPDVKPAFSSYSVQVGODLKIEVPIISGRPKPTITWTKDGLPLKQTRINVTDS	24208
QY	204	-----	203
Db	24209	LDLTLSIKETHKDDGGQYGITVANVVGQKTASIEIVTLDPDPKPGVKFDDVSAESIT	24268
QY	204	-----	203
Db	24269	LSWNPPLYTGGCQITNYIVOKRDTTITVWDVWSATVARTTLKVTKLKTGTGYQFRIFAEN	24328
QY	204	-----	203
Db	24329	RYGQSFALSDPIVAQYYPKEPGPGTFFATAISKDSMWIQWHEPVNNGSGVIGYHLER	24388
QY	204	-----	203
Db	24389	KERNISILWTKVNTIITHDTQFKAQNLKEEGIEYEFVRYAENIVGVKASKSECYVARDPC	24448
QY	204	-----	203

Db	24449	DPGTPPEIMVKNRNEITLQWTKPVYDGGSMITGYIVEKRDLPDGRWMAKASFTNVIETQFT	24508
QY	204	-----	203
Db	24509	VSGLTEDQRYEPRVIAKNAAGALSPSDSTGITAKDEVLPRI SMDPKFRDTIVVNAGE	24568
QY	204	-----	203
Db	24569	TFRLADVHGKPLPTIEMLRGDKIEIESARCEIKNTDFKALLIVKDAIRIDGGQYILRAS	24628
QY	204	-----	203
Db	24629	NVAGSKSFVPVNVKVLDRPPEGPVQVGTSEKCSLTSPPLQDGGSDISHYVVEKRET	24688
QY	204	-----	203
Db	24689	SRLAWTVASEVVTNSLKVTKLEGNEYFRIMAVNKNYGVGEPLSAPVLMKNPFVLPQP	24748
QY	204	-----	203
Db	24749	PKSLEVTNIAKDSMTVCWNRPDGSGSEITIGYIVEKRDRSGIRWIKCNKRRTDRLRLVT	24808
QY	204	-----	203
Db	24809	GLTEDHEYFRVSAENAAGVGEPSPATVYKACDPVKFPPTNAHIVDTTKNSITLAWG	24868
QY	204	-----	203
Db	24869	KPIYDGGSEILGVVVEICKADEEWOIVTPQGLRVTRFEISKLTEHQEYKIRVCALNKV	24928
QY	204	-----	203
Db	24929	GLGEATSPGTVKPEDKLEAPELDLDELKGIIVVRAGSARIHIPFKGRPTPEITWSRE	24988
QY	204	-----PEN 206	
Db	24989	EGEFTDKVQIEGVNVTQLSIDNCDNRDAGKYILKLENSGSKSAFVTVKVLDTGPGPON	25048
QY	207	-----	206
Db	25049	LAVKEVRKDSAFVWEPIIDGAKVKNVVIDKRESTRKAYANVSKSKTSFKVENLTE	25108
QY	207	-----	206
Db	25109	GAIYYFRVMAENEFVGVVETVDVAKAAEPSPGKVTLTDVDSQTSASLWMEKPEHDGG	25168
QY	207	-----	206
Db	25169	SRVLGVVEMQPKGTEKWSIVAESKVCNAVVTGLSSGQEQYQFRVKAYNEKGS DPRVLGV	25228
QY	207	-----	206
Db	25229	PVIADLTIQPSLKLPFNYSIOAGEDLKIEIPVIGRPPRNISWVKDGEPLKQTTRVNVE	25288
QY	207	-----	206
Db	25289	ETATSTVLHIKEGNKDDFGKYVTATNSAGTATENLSVIVLEKPGPPVGVPRFEVSADF	25348
QY	207	-----	206
Db	25349	VVISWEPPAYTGGCQISNYIVEKRDTTTHWMSATVARTTIKITKLKTGTEYQFRIFA	25408
QY	207	-----	206
Db	25409	ENRYGKSAPLDSKAVIVQYFFKEPGPGTPTFTSISKDOMLVQWHEPVNDGGTKIIGYHL	25468
QY	207	-----	206
Db	25469	EQEKNSILWLNKTPIODTKFTTGLDEGLEYBFKVS AENIVGIGKPSKVSCECFVARD	25528
QY	207	-----	206
Db	25529	PCDPPGRPEAIVITRNVTLKWKKPAYDGGSKITGYIVEKKDLPDGRWMAKASFTNVLETE	25588
QY	207	-----	206
Db	25589	FTVSGLVEDQRYEPRVIARNAAGNFSEPSDSSGAI TARDEIDAPNASLDPKVKDVLVHA	25648
QY	207	-----	206
Db	25649	GETFVLEADIRGKPIPDVVVMSKDGKELEETAARMEIKSTIOKTTLLVVKDCIRTDGGQYIL	25708
QY	207	-----	206
Db	25709	KLNSVGGTKSIPITVVKVLDRPPEGPEPLKVTGVTAEKCYLAWNPPLQDGGANISHYIEK	25768
QY	207	-----NYKTT-----	211
Db	25769	RETSRLSWTVQSTEVALNYSKVTLLPGNEYIFRVMVAVNKNYGIGEPLESGPVTACNPYPK	25828
QY	212	-----	211
Db	25829	PGPPSTPEVSATTKDSMVVTWARPVDDGTEIEGYILEKRDKEGVRWTKCNKKTLLDRL	25888
QY	212	-----	211
Db	25889	RVTGLTEGHSYEFRAENAAGVGEPSSEVYFACDALYPPGPPSPNPKVDTOTSRSSVSL	25948
QY	212	-----	211
Db	25949	AWSKPIYDGGAPVKYVVEVKEAADEWTTCTPTGLQKQFTVTKLKENTENFRICAI	26008
QY	212	-----	211
Db	26009	NSEGVGEPATLPGSVVAQERIEPPEIELDADLRKVVLRASATLRLFTVTKGRPEPEVKW	26068
QY	212	-----	211
Db	26069	EKAEGILTDRAGIEVTSFTMLVIDNTRFRDGRYNLTLENNSGSKTAFVNVVRVLDSPSA	26128
QY	212	-----	211
Db	26129	PVNLTIREVKKDSVTLSEWPEPLIDGGAKITNYIVEKRETRKAYATITNNCTKTTFRIEN	26188
QY	212	-----	211
Db	26189	LOEGCSYFVRVLASNEYIGLPAETTEPVKVSEPLPPGVRVTLVDVTRNTATIKWEKPS	26248
QY	212	-----	211
Db	26249	DGGSKITGYVVEWQPKGSEKWNSTCTQVKTLEATISGLTAGEEYVFRVAANKEGRSDPRQ	26308
QY	212	-----	211
Db	26309	LGVPIARDIEIKPSVELPFFHTFNVKAREQLKIDVPFKGRPOATVNRKDGOTLAKETTRV	26368
QY	212	-----	211
Db	26369	NVSSSKTVTSLSIKEASKEDVGTYELCVNSAGSITVPTIIVLDRPGLPPGPIRIDEVSC	26428
QY	212	-----	211
Db	26429	DSITISWNPPEYDGCQISNYIVEKKTSTTWHIVSOAVARTSIKIVRLTTGSEYQFRV	26488
QY	212	-----	211
Db	26489	CAENRYGKSSYSESAVVAEYFPFPPTGTPKVVHATKSTMLVTWQVPVNDGGSRVIGY	26548
QY	212	-----	211
Db	26549	HLEYKERSILWMAKANKILIAIDTOMKVSGLDEGLMYEYRVYAENIAGIKGKSKSCEPVPA	26608
QY	212	-----	211
Db	26609	RDPDPPGQPEVTNITRKSIVSLKWSKPHYDGGAKITGYIVERRELPGDRLWLCNVTNIQE	26668

Db 28829 RCNFTDVSECQYTVTGLSPGDRYEFRIARNNAVGTISPPSQSGIIIMTRDENVPPIVEFG 28888
QY 215 ----- 214
Db 28889 PEYFDGLIIKSGESLRIKALVQRPVPRVTFWKDGVIEIKRMNMEITDVLGSTSLEVRDA 28948
QY 215 ----- 214
Db 28949 TRDRHGVTVEAKNAGSAAEIKVKVQDTPGKVVGPIRFTNITGKMTLWMDAPLNDGC 29008
QY 215 -----LDSO 218
| | | |
Db 29009 APITHYIEKRETSRLAWALIEDKCAQSYTAIKLINGNEYQFRVSAVNKFGVGRPLDSO 29068
QY 219 ----- 218
Db 29069 PVVAQIQYVDPDAGIPESNITGNSITLTWARPESDGGSEIQOYLIERREKKSTRWVKV 29128
QY 219 ----- 218
Db 29129 ISKRPISETRFKVTGLTEGNEYEFHVMAENAGVGSPAGISRLIKCREPVNPPGPTVVK 29188
QY 219 ----- 218
Db 29189 VTDTSKTTVSLEWSKPVFDGMEIIGYIIMCKADIGDWHKYNAAEACVKTXYTVDLQAG 29248
QY 219 ----- 218
Db 29249 EEYKFRVSAINGAGKDSCEVTGKIKAVDRLTAPELDIDANFKQTHVVRAGASIRLFIAY 29308
QY 219 ----- 218
Db 29309 QGRPTPTAVMSKPDNSLRADIIHTDSEFTLTVCNCRNDAGKYTLTVENNNGSKSITF 29368
QY 219 ----- 218
Db 29369 TVKVLDTPGPPGPIPKDVTGRSATLMDAPLLDGGARIIHHYVVEKREASRRSQVISEK 29428
QY 219 ----- 218
Db 29429 CTRQIFKVNDAEGVPYFRVSAVNEYGVGEPEYEMPEPIVATEQPAPPRLDVDTSKSS 29488
QY 219 ----- 218
Db 29489 AVLAWLDPDGGSRITGYLLEMROKGSDFWEAGHTKOLTFTVERLVEKTEYEFVRKAK 29548
QY 219 ----- 218
Db 29549 NDAGYSEPREAFSSVIIKEPQIEPTADLTGINTQLITCKAGSPFTIDVPISGRPAKVTW 29608
QY 219 -----GSFFL----- 223
| | |
Db 29609 KLEEMRLKETDRVSITTKDRITTLTVKOSMRDGSRYFLTLENTAGVKTFSTVVVIGRP 29668
QY 224 ----- 223
Db 29669 GPTGPIEVSSVSAESCVLWSGEPKDGGEITNYIIVKRESGTTAWQLVNSSVRKTOIK 29728
QY 224 ----- 223
Db 29729 VTHLTXYMEYSFRVSSSENFVGSKPLESAPIIAEHFPVPPSAPTRPEVYHVSANAMSIRW 29788
QY 224 ----- 223
Db 29789 EEPYHDGSKIIGYWEKKERNITLWVKENKVPCLCNYKVTGLVEGLEYPQRTVALNAA 29848
QY 224 ----- 223
Db 29849 GYSKASEASRPIMQONPVDAPRPEVTDVTRSTVSLINSAPAYDGGSKVGVYIIEKKPVS 29908
QY 224 ----- 223
Db 29909 EVGDGRWLKCNITYIYSDNFTVTALSEGDTYEFVRVLAKNAAGVISKSESTGPTCRDEY 29968

QY 224 ----- 223
Db 29969 APPKAELDARLHGLDVTIRAGSDLVLDAAVGGKPEPKIITWTKGDKELDCEKVSLOYTCK 30028
QY 224 ----- 223
Db 30029 RATAVIKFCDRSDSGKYTLTVKNASCTKAVSVNVKVLDSPGCGKLTVSRTVQEKCTLAW 30088
QY 224 ----- 223
Db 30089 SLPQEDGGAETHYIVERRETSRLNMVIVEGECPTLSYVVTRLIKNNIEYIFRVRANKYG 30148
QY 224 ----- 223
Db 30149 PGVPVESEPIVARNSEFTIPSPGCIPEEVGTGKEHIIQWTKPESDGGNEISNLYVDKREK 30208
QY 224 ----- 223
Db 30209 KSLRWTRVNKYVYDTRLKVTSLMEGCDYQFRVTAVNAAGNSEPSEASNFISCREPSYT 30268
QY 224 ----- 223
Db 30269 PGPPSAPRVVDTKHSISLAWTKPMYDGTDIVGVLEMQEKDQOWYRVHTNATIRNTE 30328
QY 224 ----- 223
Db 30329 FTVPDLKMGQKYSFRVAANVVKMSEYSESAIEIEPVERIEIPDLEADDLAKTVTIRAG 30388
QY 224 -----YSKLTVDK----- 231
| | |
Db 30389 ASRLMWSVSGRPPVITWSKQIDLASRAIIDTTESYSLIIVDKVNRNDAGKYTIEAEN 30448
QY 232 ----- 231
Db 30449 QSGKKSATVLVKVYDTPGCPSPVKVKSRSVTTWETIPTIDGGAPVNNYIVKREARM 30508
QY 232 ----- 231
Db 30509 RAFKVTWTKCKTLYRISGLVEGTMYFRVLPENIYIGGPCETSDAVLSEVPLVPAKL 30568
QY 232 ----- 231
Db 30569 EYVDVTSTVTLAWEKPLYDGGSRITGYVLEACKAGTERMMKVTLKPTVLEHTVTSNE 30628
QY 232 ----- 231
Db 30629 GEQYLFIRAPQNEKGVSEPRETVTAVTVQDLRLVLPIDLSMTPOKTIHVPAGRPVELYIP 30688
QY 232 ----- 231
Db 30689 IAGRPPPAASWFFAGSKLRESERVTVETHTKVAKLTIRETTIRDTGEYTLKLNVTGTS 30748
QY 232 ----- 231
Db 30749 ETIKVIILDKPGPTGPIKIDEIDATISITISWPELDDGAPLSGVYVVEQDAHRPGWLP 30808
QY 232 ----- 231
Db 30809 VSESVTRSTFKTRLTGEGNEYFRVAATNRFGISGLQSEVIECRSSIRIPGPPELQIF 30868
QY 232 ----- 231
Db 30869 DVSRDGMTLTWYPEDDGGSQVTGYIVERKEVRADRVVRNVKVPVTMTTRYRSTGLTEGLE 30928
QY 232 ----- 231
Db 30929 YEHRTVAINARGSKPRSPKPIVAMDPIAPPKQPNPRVDTTTRTSVSLAWSVPEDEGG 30988
QY 232 ----- 231
Db 30989 SKVTGYLIEMQVQDQHEWTKCNTTPTKIREYTLTHLPOGAEYFRFRVLACNAGGPGPAEV 31048

```
QY 232 -----SRWQOG----- 237
Db 31049 PGTVKVTEMLEYPDYELDERYQEGIFVRQGGVIRLTPIKGPFPICKWTKEGODISKRA 31108
QY 238 ----- 237
Db 31109 MIATSETHTELVKEADRGDSGYDLVLENKCKKAVYIKRVIGSPNSPEGPFLYDDIQ 31168
QY 238 ----- 237
Db 31169 VRSVRVSRPPADGGADILGYILERREVPKAAWYTIDSRVGTSLVVKGLKENVEYHFR 31228
QY 238 ----- 237
Db 31229 VSAENQFISKLKSEBPVTKPLNPPPEPSPNPPELDVTKSSVLSWSRPKDDGSRV 31288
QY 238 ----- 237
Db 31289 TGYIERKETSTDKWVRHNKTQTTMTYVTGLVPDAEYQFRIQAQNDVGLSETSPASEP 31348
QY 238 ----- 237
Db 31349 VVCKDPDKPSQCELEILSISKDSVTLQWEKPECDDGKELGYWVEYRQSGDSAMKXSN 31408
QY 238 ----- 237
Db 31409 KERIKDKOFTIGLLEATEYFRFAENETGLSRPRRTAMSIKTLTSGEAPGIRKEMKD 31468
QY 238 ----- 237
Db 31469 VTTKLGBAALSCQIVGRPLDIKWYRFGKELIQSRKYKMSDGRTHLTVMTEEQDEG 31528
QY 238 ----- 237
Db 31529 VYTICATNEGEVETSSKLLQATPQHPGYPKKEYGAVGSTLRLHVMYIGRPVPAWT 31588
QY 238 -----NVF----- 240
Db 31589 WFHQKLLONSENITIENTEYTHLVNMKNVQRKTHAGKYKVLNSVFGTVDAILDVEIQD 31648
QY 241 ----- 240
Db 31649 KDKPTGVIEALLKNSAVISMKPPADGGSWITNVVKECAKEGAEWLVSSAISVT 31708
QY 241 ----- 240
Db 31709 TCRIVNLTENAGYFRVSAQNTFGISDPLEVSSVWIIKSPFEKPGAPKPTITAVTKDSC 31768
QY 241 ----- 240
Db 31769 VVAKPPASDGGAKIRNYLEKREKKONKWI SVTTEI RETVFSVKNLIEGLEVEYFRVKC 31828
QY 241 ----- 240
Db 31829 ENLGESEWSEISEPIPKSDVPIQAPHFKBELNLNVRVQSNATLVCKVTGHPKPIVKW 31888
QY 241 ----- 240
Db 31889 YROKKEIIADGLKYRIOEFKGGYHQLIIASVTDATVYQVRATNQGSGSVGTASLEVEV 31948
QY 241 ----- 240
Db 31949 PAKIHLPKTLEGMGAVHALRGEVVSIIKIPFSGKPDVITWQKQDLIDNNGHYQVIVTRS 32008
QY 241 ----- 240
Db 32009 FTSLVFPNGVERKDAGFYVVCANRFGIDQKTVELDVADVPDPGRGVKVSVDVSRDSYNLT 32068
QY 241 ----- 240
Db 32069 WTEPASDGGSKITNYIVEKCATTAERWLRVQARETRVTVINLFGKTSYOFRIAENKFG 32128
QY 241 ----- 240
```

```
Db 32129 LSKPSEPSEPTITKEDKTRAMNYDEEVDREVSMTKASHSTKELYKYEMIAEDLGRGE 32188
QY 241 ----- 240
Db 32189 FGIHRCVETSSKTYMAKFVKVKGTDQVLVKKEISILNTARHNIHLHESFESMEELY 32248
QY 241 ----- 240
Db 32249 MIFEFIGLDFERINTSAFELNEREIVSVHVQVCEALQFLHSHNIGHFDIRPENIIYQT 32308
QY 241 ----- 240
Db 32309 RRSSTIKIIEFGOARQLPGDNFRLFLTAPEYYAPEVHQHDVVSTATDMWSLGTLYVLL 32368
QY 241 ----- 240
Db 32369 SGINFLAETNQOIIENIMNAEYTFDEEAFKEISIEAMDVFDRLLVYKERSMTASEALQ 32428
QY 241 ----- 240
Db 32429 HPWLKQKIERVSTKVIRTLAHRYYHTLIKLDLNMVSAARISCGGAIRSQKGVSAKVK 32488
QY 241 ----- 240
Db 32489 VASIEIGPVSGQIMHAVGEGGHVKYCKIENYDQSTQVTWYFGVROLENSEKYEITYED 32548
QY 241 ----- 240
Db 32549 GVAILYVKDITKLDGTYRCKVVDYGEDSSYAEFLVKYGVREYDYVYCRRTMKKIKRRTD 32608
QY 241 ----- 240
Db 32609 TMRLLERPPEFTPLYNKTAIVGENVRFVGTITVHPPEPHTWYKSGOKIKPGDNKKYTF 32668
QY 241 -----SC----- 242
Db 32669 ESDGLYQLTINSVTTDDDAEYTVARNKYGEDSKAKLVTTLHPPTDSTLRPMFKRL 32728
QY 243 -----SYMHEAL----- 249
Db 32729 ANAECQBGQSVCFEIRVSGIPPTLKWCKDQPLSGLPNIEIIEHGLDYYALHIRDPLPE 32788
QY 250 ----- 249
Db 32789 DTGYRVTAINTAGSTSCQAHLQVERLYKKQEFKSEHERHVQKIDKTLRMAEILSG 32848
QY 250 ----- 249
Db 32849 TESVPLTQVAKREALAAVLYKPAVSTKVKGFEFLIEBKERKLRMPYDVPEPRKYK 32908
QY 250 ----- 249
Db 32909 QTTIEEDQRIKQFVPMDSMKWYKKIRDQYEMPGKLDRVQKRPKRIRLSRWEQFYVMLP 32968
QY 250 ----- 249
Db 32969 RITDQYRPKWRIPKLSQDDLEIVRPARRTVPSPDYDFYRPRRSIGDISDEELLPLIDD 33028
QY 250 ----- 249
Db 33029 YLAMKRTTEERLLEEBELGFSASPPSRPHFELSLRYSSPQAHVKVEETRKDFRYS 33088
QY 250 ----- 249
Db 33089 TYHIPTKAEASTSYAELRERHAQAAYRQPKORIMAREDEELLRPVTTTQHLSYKSE 33148
QY 250 ----- 249
Db 33149 LDFMSKEKSRKSRROREVTEITEIEEYVSKHAORESSSSASRLRRRRSLSPYIE 33208
QY 250 ----- 249
```

```
Db 33209 LMRPVSELIRSPQPAEEYEDDTERKSPPTERTPRSPSPVSSERSLSRFSARSARDIFS 33268
QY 250 ----- 249
Db 33269 RYESKMAALKTKTKTSERKYEVLSSQPFLLDHAPRITLMRSHRVPCGQNTREILNVQSKP 33328
QY 250 -----HN----- 251
Db 33329 TAENVKWHNGVELQESKITHYNTSGVLTLEILDCHTDGTYRAVCTNYKGEASDYATL 33388
QY 252 ----- 251
Db 33389 DVTGGDYTTASQRRDEEVRSPFPELTREAYAVSFFKKTSEMEASSSVREVKSQMTET 33448
QY 252 ----- 251
Db 33449 RESLSYHSAEMKSALEKSLKSTTRKIKTTLAARILTKPRSTVYEGESARFS 33508
QY 252 ----- 251
Db 33509 CDTDGPVPTVTLRKGVLSTARRHQVTTKYKSTFEISSVQASDEGNYSVVWENSECK 33568
QY 252 ----- 251
Db 33569 QEAFTLTQKARTEKAVTSPRVKSPBPRVKSPEAVKSPKRVKSPBPSHPKAVSPTET 33628
QY 252 ----- 251
Db 33629 KPTPEKVQHLVPSAPPKITQFLKAEKAEIAKLCVVESSLRAKEVTWYKDGKHLKEN 33688
QY 252 -----HY----- 253
Db 33689 GHFQFYSADGTVELKINLNLTSDOGYVCEISGEGTSTNLPQMGQAFKSIHEKVSKI 33748
QY 254 ----- 253
Db 33749 SETKSDQKTTSTVTRKTEPRKAPFIPSSKPVIVTGLQDTTVSSDVAKFAVKATGEPRP 33808
QY 254 ----- 253
Db 33809 TAIWTKDGAITQGGYKLSDEKGGFFLEIHKHTDTSGLYCTCVKNSAGSVSSCKLTI 33868
QY 254 -----TQKSLSPK 264
Db 33869 KAIKDEAQKSTOKTSETTPQK 33891

RESULT 14
Q10466 PRELIMINARY: PRT: 26926 AA.
AC Q10466;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DE 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Titin, heart isoform N2-B (EC 2.7.1.-) (Connectin).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HEART;
RX MEDLINE=96026330; PubMed=7569978;
RA Labelit S.; Kolmer B.;
RT "Titins: giant proteins in charge of muscle ultrastructure and
FT elasticity";
RN [2]
RP SEQUENCE OF 22277-25376 FROM N.A.
RX MEDLINE=92259380; PubMed=1582406;
RA Labelit S.; Gautel M.; Lakey A.; Trinick J.;
RT "Towards a molecular understanding of titin.";
RN [3]
RP SEQUENCE OF 1976-2014 FROM N.A.
RA Labelit S.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
[4]
RP CHARACTERIZATION.
RX MEDLINE=95331314; PubMed=7607248;
RA Gautel M.; Castiglione-Morelli M.A.; Pfuhl M.; Motta A.; Pastore A.;
RT "A calmodulin-binding sequence in the C-terminus of human cardiac
RL titin kinase.";
RL Eur. J. Biochem. 230:752-759(1995).
CC -!- FUNCTION: THIS GIANT MUSCLE PROTEIN MAY BE INVOLVED IN MUSCLE
CC ASSEMBLY AND IN MAINTAINING THE STRUCTURAL INTEGRITY OF
CC SARCOMERES. MAY HAVE PROTEIN KINASE ACTIVITY.
CC -!- ALTERNATIVE PRODUCTS: A NUMBER OF FORMS OF THIS PROTEIN ARE
CC PRODUCED BY ALTERNATIVE SPLICING WHICH DIFFER IN TISSUE
CC DISTRIBUTION. DIFFERENT SIZE TRANSCRIPTS MAY ALSO EXIST WITHIN ANY
CC ONE TISSUE. THE SEQUENCE SHOWN HERE IS THAT OF THE HEART ISOFORM
CC N2-B.
CC -!- TISSUE SPECIFICITY: MUSCLE-SPECIFIC.
CC -!- SIMILARITY: TO THE CATALYTIC DOMAINS OF OTHER SERINE/THREONINE
CC KINASES.
CC -!- SIMILARITY: BELONGS TO IMMUNOGLOBULIN SUPERFAMILY. CONTAINS 112
CC IMMUNOGLOBULIN C2-LIKE DOMAINS AND 132 FIBRONECTIN TYPE III-LIKE
CC DOMAINS.
DR EMBL; X64698; CAA45939.1; -
DR EMBL; X83270; CAA58243.1; -
DR EMBL; X64697; CAA45938.1; -
DR EMBL; X90568; CAA62188.1; -
DR EMBL; X64699; CAA45940.1; -
DR HSSP; P56276; 1TLK.
DR InterPro; IPR000282; Cytok_receptor_2.
DR InterPro; IPR000719; Buk_pkinase.
DR InterPro; IPR000577; FGGY_kin.
DR InterPro; IPR003962; FnIII_repeat.
DR InterPro; IPR003961; FnIII.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR000129; Peptidase_S24.
DR InterPro; IPR002016; Peroxidase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00041; fn3; 132.
DR Pfam; PF00047; Ig; 91.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00014; FNTYPEIII.
DR PRINTS; PR00726; LEXASERPTASE.
DR PRODOM; PD000001; Euk_pkinase; 1.
DR SMART; SM00060; FN3; 127.
DR SMART; SM00408; Ige2; 23.
DR SMART; SM00410; Ige2; 79.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00933; FGGY_KINASES_1; UNKNOWN_1.
DR PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR Muscle protein; cytoskeleton; Structural protein; Calmodulin-binding;
KW Serine/threonine-protein kinase; Alternative splicing; Repeat;
KW Immunoglobulin domain; Phosphorylation.
FT DOMAIN 1370 1389 4 X 5 AA TANDEM REPEATS OF R-M-S-P-A.
FT DOMAIN 4429 4614 GLU/LYS/PRO/VAL-RICH.
FT DOMAIN 24731 25070 CATALYTIC.
FT DOMAIN 25030 25056 CALMODULIN-BINDING.
FT MOD_RES 1372 1372 PHOSPHORYLATION (BY PDPK) (POTENTIAL).
FT MOD_RES 1377 1377 PHOSPHORYLATION (BY PDPK) (POTENTIAL).
FT MOD_RES 1382 1382 PHOSPHORYLATION (BY PDPK) (POTENTIAL).
FT MOD_RES 1387 1387 PHOSPHORYLATION (BY PDPK) (POTENTIAL).
FT MOD_RES 26171 26171 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 26178 26178 PHOSPHORYLATION (POTENTIAL).
```

```
RN [3]
RP SEQUENCE OF 1976-2014 FROM N.A.
RA Labelit S.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
[4]
RP CHARACTERIZATION.
RX MEDLINE=95331314; PubMed=7607248;
RA Gautel M.; Castiglione-Morelli M.A.; Pfuhl M.; Motta A.; Pastore A.;
RT "A calmodulin-binding sequence in the C-terminus of human cardiac
RL titin kinase.";
RL Eur. J. Biochem. 230:752-759(1995).
CC -!- FUNCTION: THIS GIANT MUSCLE PROTEIN MAY BE INVOLVED IN MUSCLE
CC ASSEMBLY AND IN MAINTAINING THE STRUCTURAL INTEGRITY OF
CC SARCOMERES. MAY HAVE PROTEIN KINASE ACTIVITY.
CC -!- ALTERNATIVE PRODUCTS: A NUMBER OF FORMS OF THIS PROTEIN ARE
CC PRODUCED BY ALTERNATIVE SPLICING WHICH DIFFER IN TISSUE
CC DISTRIBUTION. DIFFERENT SIZE TRANSCRIPTS MAY ALSO EXIST WITHIN ANY
CC ONE TISSUE. THE SEQUENCE SHOWN HERE IS THAT OF THE HEART ISOFORM
CC N2-B.
CC -!- TISSUE SPECIFICITY: MUSCLE-SPECIFIC.
CC -!- SIMILARITY: TO THE CATALYTIC DOMAINS OF OTHER SERINE/THREONINE
CC KINASES.
CC -!- SIMILARITY: BELONGS TO IMMUNOGLOBULIN SUPERFAMILY. CONTAINS 112
CC IMMUNOGLOBULIN C2-LIKE DOMAINS AND 132 FIBRONECTIN TYPE III-LIKE
CC DOMAINS.
DR EMBL; X64698; CAA45939.1; -
DR EMBL; X83270; CAA58243.1; -
DR EMBL; X64697; CAA45938.1; -
DR EMBL; X90568; CAA62188.1; -
DR EMBL; X64699; CAA45940.1; -
DR HSSP; P56276; 1TLK.
DR InterPro; IPR000282; Cytok_receptor_2.
DR InterPro; IPR000719; Buk_pkinase.
DR InterPro; IPR000577; FGGY_kin.
DR InterPro; IPR003962; FnIII_repeat.
DR InterPro; IPR003961; FnIII.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR000129; Peptidase_S24.
DR InterPro; IPR002016; Peroxidase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00041; fn3; 132.
DR Pfam; PF00047; Ig; 91.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00014; FNTYPEIII.
DR PRINTS; PR00726; LEXASERPTASE.
DR PRODOM; PD000001; Euk_pkinase; 1.
DR SMART; SM00060; FN3; 127.
DR SMART; SM00408; Ige2; 23.
DR SMART; SM00410; Ige2; 79.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00933; FGGY_KINASES_1; UNKNOWN_1.
DR PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR Muscle protein; cytoskeleton; Structural protein; Calmodulin-binding;
KW Serine/threonine-protein kinase; Alternative splicing; Repeat;
KW Immunoglobulin domain; Phosphorylation.
FT DOMAIN 1370 1389 4 X 5 AA TANDEM REPEATS OF R-M-S-P-A.
FT DOMAIN 4429 4614 GLU/LYS/PRO/VAL-RICH.
FT DOMAIN 24731 25070 CATALYTIC.
FT DOMAIN 25030 25056 CALMODULIN-BINDING.
FT MOD_RES 1372 1372 PHOSPHORYLATION (BY PDPK) (POTENTIAL).
FT MOD_RES 1377 1377 PHOSPHORYLATION (BY PDPK) (POTENTIAL).
FT MOD_RES 1382 1382 PHOSPHORYLATION (BY PDPK) (POTENTIAL).
FT MOD_RES 1387 1387 PHOSPHORYLATION (BY PDPK) (POTENTIAL).
FT MOD_RES 26171 26171 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 26178 26178 PHOSPHORYLATION (POTENTIAL).
```

```
FT MOD_RES 26184 26184 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 26190 26190 PHOSPHORYLATION (POTENTIAL).
FT CONFLICT 22277 22277 T -> P (IN REF. 2).
FT CONFLICT 22449 22449 E -> G (IN REF. 2).
FT CONFLICT 22454 22454 T -> Q (IN REF. 2).
FT CONFLICT 23324 23324 S -> L (IN REF. 2).
SQ SEQUENCE 26926 AA; 2993428 MW; D5EECD3254DF5523 CRC64;

Query Match      33.9%; Score 485; DB 4; Length 26926;
Best Local Similarity 0.7%; Pred. No. 0.0002;
Matches 177; Conservative 44; Mismatches 42; Indels 23713; Gaps 55;

QY 2 WLK----- 4
   |||
Db 1543 WLKNSDIIIVPHKPIRIBGTGKAALKIDSTVSQDSAWYATATNAKAGRDTTRCKVNVYE 1602
QY 5 -----AFYDK----- 9
   |||
Db 1603 VFAEPEPERKLIIPRGTYRAKEIAAPELEPLHLRYGQEWEEGLDYDKEKQKPFPPKK 1662
QY 10 ----- 9
Db 1663 LLSRLRKGPAHFECRLTPISDPTMVVEWLHDGKPLEANRLRMINEFGYCSLDYGVA 1722
QY 10 ----- 9
Db 1723 SRDSGIITCRATNKYGTDHTSATLIVKDEKSLVEESQLPEGRKGLQRIEELERMAHEGAL 1782
QY 10 ----- 9
Db 1783 TGVTTDQEKQKPDIVLYPEPVRLVEGETARFCRVGTGYPQPKVNWYLNGLIRSKRFR 1842
QY 10 ----- 9
Db 1843 VRYDGIHYLDIYDCKSYDTGEVKTVAENPEGVIEHKVLEIQOOREDFRSVLRAPEPRPE 1902
QY 10 ----- 9
Db 1903 FHWHPGKLQFEVQKVRDPVDTTETKEVVKLRAERITHEKVPESSEELRSFKRRTEBG 1962
QY 10 ----- 9
Db 1963 YYEATAVELKSRKKDESYEELLRTKDELHWTKELTEEEKKALAECKIITPTFKPKD 2022
QY 10 ----- 9
Db 2023 IELSPSMEAPKIFERIQSQTVGQSDAHPFRVRVVGKPDPECEWYKNGVKIERSDRIYWTW 2082
QY 10 ----- 9
Db 2083 PEDNVCELVRDVTAEADSASIMVKAINIAGETSSHAFLVLQAKLITFTQELQDVVAKK 2142
QY 10 ----- 9
Db 2143 DTWATFECTSEPFVKVWKYKDGMEVHEGDKYRMHSDRKVHFLSLITIDTSDAEDYSCVL 2202
QY 10 ----- 9
Db 2203 VEDENKVTAKTLIVEGAVVEFKELQDIEVPESYSSELECIIVSPENIEGKYHNDVELKS 2262
QY 10 ----- 9
Db 2263 NGKYTTISRGRQNLTVKDVTKEDQGEYSFVIDGKTKTCKLKMKPRIAILQGLSDQKVC 2322
QY 10 ----- 9
Db 2323 EGDIVOLEVKVLSVESVEGVMKQGEVQPSDRVHIIVIDKQSHMLLIEDMTKEDAGNYSPT 2382
QY 10 ----- 9
Db 2383 IPALGLSTGRYSVSDVITPLKDVNVIEGTKAVLECKVSPDVTSVKMWLNDEQIKPD 2442
QY 10 ----- 9
```

```
Db 2443 DRVQAIVKGTQKQRLVINRTHASDEGPYKLIVGRVETNCNLNLSVEKIKIIRGLRDLCTETQ 2502
QY 10 ----- 9
Db 2503 NVVFEVELSHSGIDVLWNFKDKXKIKPSSKYKIEAHGKIYKLVNLMMKDDCKGYTFYAGE 2562
QY 10 ----- 9
Db 2563 NMTSGKLIVAGGAISKPLTDQIVAESQEAFFEVEANPDPSKGEWLDRGKHLPLTNNIRSE 2622
QY 10 ----- 9
Db 2623 SDGHRRLIIAATKLDIDIGEYTYKVATSKTSAKLVEAVKIKTKLNLTVTTQDAVFTV 2682
QY 10 ----- 9
Db 2683 ELTHPNVGVQVQKNGVVLÉSNEKYAISVKGTYISLRINKCAIVDESIVYGFRLGLGASA 2742
QY 10 ----- 9
Db 2743 RLHVETVKIIKPKDVTALENATVAFEVSVDHTVPVWFKHSVEIKPSDKHRLYSERKV 2802
QY 10 ----- 9
Db 2803 HKLMLQNISPDAGEYTAVVGOLECKAKLFVETLHITTKMKNIEVPETKTASFCEVSHF 2862
QY 10 ----- 9
Db 2863 NVPSMWLKNVGEIEMSEKFIIVQGLHOLIIMNTSTEDSAEYTFVCGNDQVSATLTVTP 2922
QY 10 ----- 9
Db 2923 IMITSMKDINAEEKDTITFEVTNVYEGISYKWLKNGVEIKSTDCKQMTKLTHTSLNIR 2982
QY 10 ----- 9
Db 2983 NVHFGDAADYTFVAGKATSTATLYVEARHIEFRKHIDIKVLEKKRAMFECEVSEPDITV 3042
QY 10 ----- 9
Db 3043 QMKDDQELQIDTRIKIQEKYVHRLILIPSTRMSDAGKYTVVAGGNVSTAKLFEVGRDVR 3102
QY 10 ----- 9
Db 3103 IRSIKKEVQVIEKQRAVVEFEVNEDDVDAHVKDGIIEINFOYQERHKYVVERRIHRMFTS 3162
QY 10 ----- 9
Db 3163 ETRQSDAGEYTFVAGNRSSVTLYVNAPEPQVLOLQPVTVQSGKPARFCAMISGRPOP 3222
QY 10 ----- 9
Db 3223 KISWYKEBOLLSTGFKCKFLHDGQEYTLILLIEAFPEDAAVYTCEAKNDYGVATTASLSV 3282
QY 10 ----- 9
Db 3283 EYVEVSPDQEMPVYPAIITPLQDVTVTSEGQPARFCQVSGTDLKVSWSYKDKKIKPSR 3342
QY 10 ----- 9
Db 3343 FFRMTQFEDTYQLEIAEAYPEDEGTYTFVANNAGVSSSTANLSLEAPESILHERIEQEI 3402
QY 10 ----- 9
Db 3403 EMEMKEFSSFLSAEEGLHSAELQLSKINETLELLESFVPTPKFDSKEGTGPIFIKE 3462
QY 10 ----- 9
Db 3463 VSNADISMGDVATLSVTVIGIPKPKIQWFFNGVLLTPSADYKVFVDDGDDHSLIILFTKLE 3522
QY 10 ----- 9
```

Db 3523 DEGEYTCMASNDYKGTICSAYLKINSKGEGHODTETESAVALSKLEKLGPGCPHFLKELK 3582
Qy 10 ----- 9
Db 3583 PIRCAOGLPAIFEYTVVGEPAFTVTFWKENQOLCTSVYTTIHNPNNGSGTFIVNDPQRED 3642
Qy 10 ----- 9
Db 3643 SGLYICKAENLGESTCAAEALLVLEDDTMDTPTCKAKASTPEAPDFPQTPKGPAAVEAL 3702
Qy 10 ----- 9
Db 3703 DSEQEIATFVKDTILKAALITEENOQLSYEHIAKANELSSOLPLGAOELQILEQDKLTP 3762
Qy 10 ----- 9
Db 3763 ESTREFLCINGSIHFOPLKEPSNLOLQIVOSQKTFESKEGILMPPEPEPTQAVLSDTEKIF 3822
Qy 10 -----VAE----- 12
Db 3823 PSAMSIQINSLTVEPLKTLAEPEGNYPQSSIBPPMHSYLTSAEVLSLKERTVSDTN 3882
Qy 13 ----- 12
Db 3883 REQVRTLOKQEAQASALILSQAEGHVESLQSPDVMISQVNYEPLVPSEHSCTEGGKILI 3942
Qy 13 ----- 12
Db 3943 ESANPLENAGODSAVRIEKGSLRFPLALEEKQVLLKEHSDNVVMPDQIIIESKREPA 4002
Qy 13 ----- 12
Db 4003 IKKVOGRODLLSKESLLSGIPBEORNLKIQICRALQAASEQPLGFSEWLRNTEKV 4062
Qy 13 ----- 12
Db 4063 EVEAVNITQEPRHIMCWLTVTSKASVTEVTIIEDVDPQMANLKMELRDALCAIIEEI 4122
Qy 13 ----- 12
Db 4123 DILTAEGPRIOGAKTSLQEEMDSFGSQKVEPITEPEVESKYLISTEVSFYFNQSRVK 4182
Qy 13 -----KLKEA----- 17
Db 4183 YLDATPVTKGVASAVSDEKQDESILKPEKEESSESSEGTBEVATVKIQEAEGGLIKEDG 4242
Qy 18 ----- 17
Db 4243 PMIHTPLVDTVSEGDIVHLTTSITNAKEVNWYFENKLVPSDEKFKCLQDQNTVTLVIDK 4302
Qy 18 ----- 17
Db 4303 VNTEDHQGEYVCEALNSGKTATSAKLTVVYKRAAPVIKRKTEPLEVALGHLAKETCEIQS 4362
Qy 18 -----FDWLKA----- 23
Db 4363 APNVRFQWFKAGREIYESDKCSIRSSKYISSLEILRTQVVDGEYCTCKASNEYSVSCTA 4422
Qy 24 ----- 23
Db 4423 TLTVTVPGGKKVRKLLPERKPEPKPEVVLKSVLKRPEEPEPKLEKVKKPAVPE 4482
Qy 24 ----- 23
Db 4483 PPPKPVVEEVPVTKRERKIPEPTKVPEIKPAIPLPAPEPKPKPEAEVKTIKPPVP 4542
Qy 24 ----- 23
Db 4543 EPTPIAAPVTVPVVGKRAEAKAPKEAAKPGPIKGVPKTPSPIEARRKLRPGSGEK 4602
Qy 24 ----- 23
Db 4603 PPDEAPPTYQLKAVPLKFVKEIKDIILTESEFVGSSAIFECLVSPSTAITTTMMKDGSNIR 4662

Qy 24 ----- 23
Db 4663 ESPKXRFIADGKDRKLHIIDVQLSDAGEYTCVLRGNKKEKSTAKLVVEELPVRFVKTLK 4722
Qy 24 ----- 23
Db 4723 BEVTVVRGQPLYLSCLENKERDVVWRKDKIVVEKPGRIVPGVIGLMLRALTINDADDTDA 4782
Qy 24 ----- 23
Db 4783 GTTYTIVTENANNLECCSCVKVVEVIRDWLVKPIRDQHVKPGTAFACDTAKDTPNKKWF 4842
Qy 24 ----- 23
Db 4843 KGYDEIPAEPNDKTEILRDGNHLYLKIKNAMPEDIAEVAVEICKRYPAKLTLCEREVEL 4902
Qy 24 ----- 23
Db 4903 LKPIEDVTIYKESASFDAEISEADIPQOMWKLKCELLRSPSTCEIKAEKKRFLTLHKVK 4962
Qy 24 ----- 23
Db 4963 LDQAGEVLYQALNAITTAILTVEIELDFAVPLKDVTPPERQARFECVLTREANVWSK 5022
Qy 24 ----- 23
Db 5023 GPDIIKSSDKFDIADGKKHILVINDSQFDDDEGVYTAEEVGKKTSAFLVFTGIRLKFMS 5082
Qy 24 ----- 23
Db 5083 LEDQTVKEGETATFVCELSHEKMHVWFKNDAKLHTSRTVLISSEKTHKLEMKEVTLDD 5142
Qy 24 -----FYDKVAE----- 30
Db 5143 ISQIKAOVKELSSAQKLVLEADPYFTVKLHDKTAVEKDEITLKEVSKOVVWKFWDGE 5202
Qy 31 ----- 30
Db 5203 EIVSPKYSIKADGLRILKIKADLKGEYVDCGDKTKANVTVEARLIEVEKPLYG 5262
Qy 31 ----- 30
Db 5263 VEFVGETAFHEIELSEPDVHGOMWKLKGOPLTASPDCEIIEIDGKKHILHNCOLGTMTE 5322
Qy 31 -----KLKE----- 34
Db 5323 VSPQANAKSAANLKVAKELPLIFITPLSDVKVFEKDEAKFECEVSRPEKTRWLGKTOEI 5382
Qy 35 -----AFMD----- 38
Db 5383 TGDREFELIKDGTKHSWIKSAFEDEAKYFMAEDKHTSGKLLIEGIRLKLFTPLKDV 5442
Qy 39 ----- 38
Db 5443 AKESAVFTVELSHDNIRVKFNQDLHTTRSVSMQDEKTHSITFKDLSIDDTSQIR 5502
Qy 39 ----- 38
Db 5503 VEAMGMSSEAKLTVLEGDYFTGKLQDYTGVEKDEVILQCEISKADAPVWFKDGEIKP 5562
Qy 39 ----- 38
Db 5563 SKNAVINTDGKRRMLILKALKSDIGQYTCDCGDTKTSGLDIEDREIKLVRPLHSVEVM 5622
Qy 39 ----- 38
Db 5623 ETETARFETISEDIIHANWKLKGEALLQTPDCEIKEEGKIHSILVHNCRLDQTTGGVDFQ 5682
Qy 39 ----- 38
Db 5683 AANVKSALHVRVPRVTGLLRPLKDVTVTAGETATFDCELSYEDIPVEWYLGKKLEPSD 5742

QY 39 ----- 38
Db 5743 KVPSEGVHTLTLRDVKLEDAGEVQLTAKDFKTHANLFVKPEPVEFTKPLEDQTVBEG 5802
QY 39 -----KTHTC--- 43
Db 5803 ATAVLECEVSRENKVKFKNGTEILKSKYIEIVADGRVKLVTHDCTPEDIKTYTCDAK 5862
QY 44 ----- 43
Db 5863 DEKTSCLNVPPHVEFLRPLTDLOVREKEMARFCELSRENKVKWFKDGAIEKKKKY 5922
QY 44 ----- 43
Db 5923 DIISKAVRILVINKCLLDEAEYSCEVRTARTSGLTVLEEAFTKLANIEVSETDT 5982
QY 44 ----- 43
Db 5983 IKLCEVSKPGAETWYKGBDEIIEGTGRYIELTEGRKKRILVIONAHLNEDAGNYNCRLPSS 6042
QY 44 ----- 43
Db 6043 RTDGVKVKHELAAEFISKPNLEILLEGEKAFFVCSISKESPPVQWKRDDKTLES GDKYDV 6102
QY 44 ----- 43
Db 6103 IADGKRKVLVVKDATLQDMGTYYVMVGAARAAHLTVIEKLRIVVPLKDRVKEQEVVF 6162
QY 44 ----- 43
Db 6163 NCEVNTGAKAKWFNEBAIFDSSKYIILOKDLVYTLIRDAHLDDQANVNSLNRHGE 6222
QY 44 ----- 43
Db 6223 NVKSAANLIVEEDLRIVEPLKDIETMEKKSFTFWCKVYNRLNVLTKWTKNGEEVFPDNRV 6282
QY 44 ----- 43
Db 6283 SYRVDKYKHLTIKDCGFPDEGEYIVTAGODKSVAELLIIEAPTEFVEHLEDQTVTERDD 6342
QY 44 ----- 43
Db 6343 AVFSCQLSRENKVKWYRNGREIKGKKYKEKGSIHRLIIKDCRLDDECEYACGVEDR 6402
QY 44 ----- 43
Db 6403 KSRARLFVEEIPVEIIRPPQDILEAPGADVFLAELNKDKVEVQWLRNNWVYVQDKHQM 6462
QY 44 ----- 43
Db 6463 MSEGKIHLQICDIKPRDQGEYRFLAKDKEARAKLELAAAPKIKTADQDLVVDVGKPLTM 6522
QY 44 ----- 43
Db 6523 VVPYDAYPKAEAEWFKENEPLSTKTIDTAQTSFRILEAKKGDKGRYKIVLQNHKGAE 6582
QY 44 ----- 43
Db 6583 GFINKLVIDVPGVRNLEVTFETFDGEVSLAWEEPLTDGSGKIIGYVVERDRIKRTWVLA 6642
QY 44 ----- 43
Db 6643 TDRACESFTVTGLOKGGVEYLFVRSAARNVGTGEPVETDNFVEARSKYDVPGLNVTI 6702
QY 44 ----- 43
Db 6703 TDVNRFGVSLTWEPPEYDGAETWYIELRDKTSIRMDTAMTVRAEDLSATVTDVVEGQ 6762
QY 44 ----- 43
Db 6763 EYSFRVRAQNRIGVGKPSAATPFVKVADPIERPSPVNLTSDDQTSQSVOLKWEPLKDG 6822
QY 44 ----- 43

Db 6823 GSPILGYIIERCEGKNWIRCNMKLVLPETYKVTGLEKGNKYLXRVSAENKAGVSDPSE 6882
QY 44 ----- 43
Db 6883 ILGPLTADDAFVEPTMDLSAFKDGLEIVVPNPITILVPSTGYPRPTATWCFGDKVLETCGD 6942
QY 44 ----- 43
Db 6943 RVKMKTLNAYAEVLSPERSDKGIYTLKLENVKTISGEIDVNVNIARPSAPKELKFGDI 7002
QY 44 ----- 43
Db 7003 TKDSVHLTWEPDDDDGSGPLTGYVVEKREVSRTWTVMDFVTDLEFTVPDLVOGKEYLF 7062
QY 44 ----- 45
Db 7063 KVCARNKCGPEPAYVDEPVNMSTPATVPDPNPNVKNRDRTANSIFLTWDPKNDGGSRI 7122
QY 46 ----- 47
Db 7123 KGYIVERCPRGSDKWKVACGEPVAETKMEVTGLEEGKWYAYRVKTLNRQGASKSPRTEI 7182
QY 48 ----- 52
Db 7183 QAVDTQEAPEIFLDVKLLAGLTVKAGTKIELPATVTGKPEPKITWTKADMILKQDKRITI 7242
QY 53 ----- 58
Db 7243 ENVPKSTVTIVDSKRSRTGTIIIEAVNVCGRATAVVEVNVLDKPGPPAFAFDITDVTNES 7302
QY 59 ----- 58
Db 7303 CLLTNWPPRDDGSGKITNVVVERRATDSRVHKLSTVKDTNFKATKLIPKEYIFRVAA 7362
QY 59 ----- 58
Db 7363 ENMYGAGEPVOASITAKYQFDPGPPPTRELPSTDITKDAVTLTWCEPDDGSGPITGVV 7422
QY 59 ----- 58
Db 7423 ERLDPTDKWRCNMPVKDITYRVKGLTNKKYRFRVLAENLAGPKPSKSTEPILLIKD 7482
QY 59 ----- 58
Db 7483 PIDPPWPGKPTVKDVGKTSVRLNWTKPEHGGAKIESVIEMLKTGTDEWVRVAGVPT 7542
QY 59 ----- 64
Db 7543 TQHLPLGLMEGOEYSFRVRAVNKAGESEPSDPVLCREKLYPPSPRWLEVINITKNT 7602
QY 65 ----- 73
Db 7603 ADLKWTVPKGGSPITNIVIVEKRDVRRKGWOTVDTTVKDTCTVTPTLREGSLYFRVAA 7662
QY 74 ----- 83
Db 7663 ENAIGQSDYTEIEDSVLAKDTFTTGPYPVALAVDVTKRHDVKWEPKNDGGRPIQRVV 7722
QY 84 ----- 86
Db 7723 IEKKERLGRWYKAGTAGPCNFRVTDVIEGTVOFVRAENEAGVGHSPSEPTILSIE 7782
QY 87 ----- 88
Db 7783 DPTSPSPPLDLHVTDAGRKHIAIAWKPPKNGGSPIIIGYHVEMCPVGTEKMWVRNRPDI 7842
QY 89 ----- 92
Db 7843 KDLKFKVEGVVDPKKEYVLRVRAVNAIGYSEPSISENVVAKDPCKPTIDLETHDIIVI 7902
QY 93 ----- 98
NWYVDG
: : : :

Db	7903	EGEKLSIPVPPRAVPVPTVSHHKGKVKASDRLTMKNDHISAHLEVPKSVRADAGIYTI	7962
Qy	99	-----	98
Db	7963	TLENKLSATASINVKVIGLPCPKDIKASDITKSSCKLTWEPPEFDGTPILHVLERR	8022
Qy	99	-----	98
Db	8023	EAGRTYIPVMSGENKLSWTVKDLIPNGEYFFRVKAVNKVGGEYIELKNPVIADPKQP	8082
Qy	99	-----	104
		VEVHNA-----	
Db	8083	PDPPVDVEVHNPTAEAMTITWKPPPLYDGGKIMGYIEKIAKEERWKRCHLVPILTY	8142
Qy	105	-----	104
Db	8143	TAKGLEEKEXQFVRVRAENAAGISEPSRATPPTKAVIDPDKVILRTSLEVKRGDEIAL	8202
Qy	105	-----	104
Db	8203	DASISGSPYTTITWIKDENIVPEIEKKRAAPLVRRRKGEVQEEPPVPLTLQRLSIDNS	8262
Qy	105	-----	104
Db	8263	KGESQLVRDSLRDPDHGLYMIKVENDHGIAPCTVSVLDTGPPINFVPEDIRKTSVL	8322
Qy	105	-----	108
		-----KTKP-----	
Db	8323	CKWEPPLDGSGEINIYLEKKDKTKPDSEWIVVTSILRHCKYSVTKLIEGKEYLFRVA	8382
Qy	109	-----	108
Db	8383	ENRFGPGPCVSKPLVAKDPFGPDAPDKPIVEDVTSNSMLVKWNEPKDNGSPILGYWLE	8442
Qy	109	-----	108
Db	8443	KREVNSTHWSRVNKSLLNALKANVDLLEGLTYFVRCAENAAAGPGRFPPSDPKTAHP	8502
Qy	109	-----	108
Db	8503	ISPPGPPIPRYDTSITIELEWEPFANGGEIVGVFDKQLVGTNKNWSRCTEKMIVR	8562
Qy	109	-----	108
Db	8563	QYTVKEIREGADYKLRVAVNAAGEGPPGETQPTVAEPQPPAVELDVSVKGGIIMAG	8622
Qy	109	-----	108
Db	8623	KTLRIPAVVTGRPVTKVWTKGEGELOKDRVVIDNVGTKSELIITKDALRKDHGRYVITAT	8682
Qy	109	-----	108
Db	8683	NSCGSKFAAARVEVDVPGVLDLKPVTYNRKMCLLNSDPEDDGGSEITGFIIERKDAK	8742
Qy	109	-----	108
Db	8743	MHTWRQPIETERSKCDITGLGEOGEYKFRVIAKNKFGCGPVEIGPILAVDPLGPPTSPE	8802
Qy	109	-----	108
Db	8803	RLTYTERQRSTITLDWKEPRNSGSPIOGYIIIEKRRHDKPDFRVNKRCLCTTSFLVENL	8862
Qy	109	-----	108
Db	8863	DEHQMYEFRKAVNAIGESEPSPLVNVVQDDVEPPTIKLRLSVRGDTIKVKAGEPVHIP	8922
Qy	109	-----	108
Db	8923	ADVTGLPMPKIEWSKNETVIEKPTDALQITKEEVSERSEAKTELSIPKAVREDKGTYTVA	8982
Qy	109	-----	108
Db	8983	SNRLGSVFRNVHVEYDRPSPRNLAVTDIKAESCYLTDWAPLDNGSGSEITHYVIDKRDA	9042

Qy	109	-----	108
Db	9043	SRKKAWEVINTAVEKRYGIWKLIPLNGOYEFRRVAVNKYGISDECKSDKVVIQDPRLP	9102
Qy	109	-----	108
Db	9103	GPPGPKVLARTKGSMLVSWTPPLDNGGSPITGYWLEKREEGSPYWSRVSRAPITKVGLK	9162
Qy	109	-----	108
Db	9163	GVEFNVPLLEGVYQFRAMAINAAGIGPPSEPSDEVAGDPIFPPGPPSCEPVKDKTKS	9222
Qy	109	-----	108
Db	9223	SISLGWPPAKGGSPIKGYIVEMQEECTTDWKRNVNEPDKLITTCCECVVNLKELRKYRF	9282
Qy	109	-----	108
Db	9283	RVKAVNEAGESEPSDTTGEIPATDIOEBPEVFDIGAQDCLVCKAGSQIRIPAVIKGRPT	9342
Qy	109	-----	108
Db	9343	PKSWEFDGKAKAMKGVHDIPEDAQLETAENSSVIIIECKRSHTKYSITAKNKAGQ	9402
Qy	109	-----	108
Db	9403	KTANCRVKVMDVPKPKDLKVSDITRGSCRLSWKMPDDDDGGDRINKGYVIEKRTIDGKAWT	9462
Qy	109	-----	113
		-----REEQY-----	
Db	9463	KVNPDCGSTTFWPDLLSEQQYFFRVAENRFGIGPPVETIQTARTARDPIYPPDPPIKLK	9522
Qy	114	-----	113
Db	9523	IGLITKNTVHLSWKPPKNDGGSPVTHYVECLAWDPTGTKEAWRQCNKRDEVELOFTVE	9582
Qy	114	-----	113
Db	9583	DLVEGGEYFVRKAVNAAGVSKSATVGPCQCRDMPPSIDLKEFNEVEEGTWNVIVAK	9642
Qy	114	-----	113
Db	9643	IKGVFPPTLTFKAPKKPNKPNKPELVYDTHVNKLWVDDTCTLVIPQSRSDTGLYTITAV	9702
Qy	114	-----	113
Db	9703	NNLTASKEMRLNVLGRGPPVGPPIKFESVSADQMTLSWFFPKDDGGSKITNYVIEREA	9762
Qy	114	-----	113
Db	9763	NRKTWVHVSSEPKECTYTIPLKLEGEYVFRIMAQNKYIGICEPLDSBPETARNLFSVPGA	9822
Qy	114	-----	113
Db	9823	PKPTVSSVTRNSMTVNWEEPEYDGGSPVTGYWLEMKDITTSKRWNKRVNRDPKAMTLGVS	9882
Qy	114	-----	113
Db	9883	YKVTGLIEGSDYQPRVYAINAAGVPASLPSPDATARDPIAPGPPPKVTDWTKSSADL	9942
Qy	114	-----	113
Db	9943	EWSPPLKGGSKVTGYIVEYKEEGKEBEWKGKKEVRGRTKLVTGLKEGAFYKFRVSAVN	10002
Qy	114	-----	113
Db	10003	IAGIGEPGEVTDVTEMKDRLVSPDLQDASVDRIVVHAGVIRIIAYVSGKPPPTWTN	10062
Qy	114	-----	113
Db	10063	MNERTLPOEATIIETAISSSMWIKNCORSQHGVSLLAKNAGEKKKTIIVDVLDPGPV	10122

QY 114 ----- 113
Db 10123 GTPFLAHLNTNESCKLTWFSPEDDGGSPITNVYIEKRESRRRAWTPVTVTVTRQNAVQ 10182
QY 114 ----- 113
Db 10183 LIQKAYFFRIAENSIGMPFVETSEALVIREPITVPERPELVEKVTNKTVLTWNP 10242
QY 114 ----- 113
Db 10243 PRYDGSSEIINYLESRLIGTEKFHKVTNDNLLSRKYTVKLGKEDTYEYRVSANNVQ 10302
QY 114 ----- 113
Db 10303 GRPSFCTKPTCKDELAPPTLHLDFRDKLTIRVGEAFALTGRYSGRKPCKVSWFKDEADW 10362
QY 114 -----NST----- 116
Db 10363 LEDDRTHIKYTPATLALEKIKAKRSDSGKYCVVVENSTGSRKGCQVNVVDHPGPPGV 10422
QY 117 ----- 116
Db 10423 SFDEVTKDYWISWKPLDDGGSKITNYIIEKKEGVQVMPVTSASAKTTCVKSKLLE 10482
QY 117 ----- 116
Db 10483 KDIYIFRIHAENLYGSDPLVSDSMKAKDRFRVPDAPQPIVTEVTKDSALVTWPKPHDGG 10542
QY 117 ----- 116
Db 10543 KPITNYILEKRETMKRWARVTKDPIHPYTKFRVPDLEGCQYFRVSAENEIGDPS 10602
QY 117 ----- 116
Db 10603 PSKPVFAKDPIAKSPVNPALDITCNSVDLTWQPRHDGSKILGIYIEYQVKVQDEW 10662
QY 117 -----YRVSV----- 122
Db 10663 RRANHTPESCETKYKVTGLRDGQYKFRVLAVNAAGESDPAHPVPVLVKORLEPELI 10722
QY 123 ----- 122
Db 10723 LDANMAREOHKIVGDTLRLSAILIKGVFPKVTWKEDRDAPTAKRIDVTPVGSKLEIRNA 10782
QY 123 -----LTV----- 125
Db 10783 AHEDGGIYSLTVENPAGSKTVSKVLVLDKPGPRDLEVSEIRKDCSYLTWKEPLDDGS 10842
QY 126 ----- 125
Db 10843 VITNYVVERRDVASQWSPLSATSKKSHFAKHLNEGNOYLFRAVAENOYGRGPFVETPK 10902
QY 126 -----LH-----ODWL-- 131
Db 10903 PIKALDPLHPGPPKDLHVDVDKTEVSLVWNKPDGSGSPITGYLVEYQEBGTQDWIKF 10962
QY 132 ----- 131
Db 10963 KVTNLECVVTGLQOGKTYFRVRAENIVGLGLPDTTPIECQEKLVPPSVELDVKLIG 11022
QY 132 ----- 131
Db 11023 LVVKAGTVRFPAILIRGVPTAKWTDTGSEIKTDEHYTVETDNFSSVLTIKNCLRRDTG 11082
QY 132 ----- 131
Db 11083 EYQIIVSNAAGSKTVAVHLTVLDVPGPTGPINILDVTPHEMTISWQPKDGGSPVINY 11142
QY 132 ----- 131
Db 11143 IVEKODTRKDTWGVSSGSSKTKLIPHLOKCEYFRVRAENKIGVGPPLDSTPTVAKH 11202
QY 132 ----- 131

Db 11203 KESPPSPCKPVVTDITENAAVTSWLTPKSDGSGPITGYMERREVTGKVRVKNKPIAD 11262
QY 132 ----- 131
Db 11263 LKFRVTGLYEGNTYFRVFAENLAGLSKPSSPDPIKACRPKIPKPPGPPINPKLKDSRET 11322
QY 132 ----- 131
Db 11323 ADLVWTKPLSDGSGPILGYVVECOQPGTAQWNRINKDELIRQCAFVRPGLIEGNEYFRRI 11382
QY 132 ----- 131
Db 11383 KAANIVGEGEPRELAESVIAKDLHPPEVELDVTCDVITVRVGTIRILARVKGREPDP 11442
QY 132 ----- 131
Db 11443 ITWTKEGKVLVREKRVLDLQDLPRVELQIKEAVRADHGKYIISAKNSSGHAQSAIVNVL 11502
QY 132 ----- 131
Db 11503 DRPGPCQNLKVTNVTKENCTISWENPLDNGGSEITNFIVEYRKPNOGWSIVASDVTKRL 11562
QY 132 ----- 131
Db 11563 IKANLLANNEYFRVCAENKVGVGPTIETKTPILAINPIDRPGEPENLHIADKGTFFVL 11622
QY 132 ----- 131
Db 11623 KWRRPYDGGSPNL SYHVERRLKGSDDERVHKSGIKETHYMWVDRVCENOIYEFVRVOTKN 11682
QY 132 ----- 131
Db 11683 EGGESDWKTEEVVVKEDLQKPVLDLKLGLVTKAGDTIRLEAGYRGKPFPEVAWTKDK 11742
QY 132 ----- 131
Db 11743 DATDLTRSPRVKIDTRADSSKFSLTRAKRSDGGKYVVTATNTAGSEFAYATVAVNLDPKPG 11802
QY 132 ----- 131
Db 11803 VRNLKIVDVSSDRCTVCWDPPEDDGGCEIOTNYILECETKRMVWSTYSATVLTPTGTVTR 11862
QY 132 ----- 131
Db 11863 LIEGNEYIFRVAENKIGTGPPTESKPVIAKTKYDKRGPDPPEVTKVSKEEMTVVWNP 11922
QY 132 ----- 131
Db 11923 EYDGGKSITGYFLEKKEKHSTRWVPVYNKSAIPERRMKVONLLPDHEYQFRVKAENEIGIG 11982
QY 132 ----- 131
Db 11983 EPSLSPRPVAKDPIEPPGPTNFRVVDTTKHSITLGMGKPVYDGGAPIIGYVEMRPKI 12042
QY 132 ----- 131
Db 12043 ADASPDGKWKRCNAAAQLVRKEFTVTSLDENOEYEFVCAQNOVGIGRPAELKEAIKPRE 12102
QY 132 ----- 131
Db 12103 ILEPPEIDLDSMRKLVIVRAGCPIRLFAIVRGPAPKVTWRKVGDVNVVRKGOVDLVD 12162
QY 132 ----- 131
Db 12163 MAFLVIPNSTRDSDSGKYSLTLVNPAGEKAVFVNVRLDTPGPVSDLKVSVDVTKTSCHVSW 12222
QY 132 ----- 131
Db 12223 APPENDGGSQVTHYIVEKREADRKTWSTVTPVEVKTSFHVNTLVPCNEYFRVTAVNEYG 12282
QY 132 ----- 131

Db 12283 PGVPTDVPKPVLASDPLSEDPDRKLEATEMTKNSATLAWLPPLRGGAKIDGYIISYRE 12342
Qy 132 ----- 131
Db 12343 EQPADRWTEYSVKWDLVLVVTGLKEGKKYFRVAARNAGVSLPREAGVYEAKEQLP 12402
Qy 132 ----- 131
Db 12403 PKILMEQITIKAGKLRLEAHVYCKPHTCKWKKGEDENVYTSSHLAVHKADSSILIK 12462
Qy 132 ----- 131
Db 12463 DVTRKDSGYSLTAENSSGTDQIKIKVVMDAPGPPQPFDISDADACSLSHWPLED 12522
Qy 132 ----- 131
Db 12523 GGSNTNIVEKDVSRGOWTALASVTKTSCRVGKLLPGQYIFRVAENRFGISEPLT 12582
Qy 132 ----- 131
Db 12583 SPKMAQFPFGVPSEPKNARTKYNKDCIFVANDRPDSGGSPILIGYLLIERKERNLLWV 12642
Qy 132 ----- 131
Db 12643 KANDTLVRSTYPCAGLVEGLEYSFRIYALNKAGSSPPKTEYVTARMPVDPGKPEVI 12702
Qy 132 ----- 131
Db 12703 DVTKSTVLIWARKHGGSKLIIGFVEACKLPDGKWRONTAPHQIPQOEYATGLEEK 12762
Qy 132 ----- 131
Db 12763 AOYQRAIARTAVNISPPSPDPVTILAENVPPRIDLSVAMKSLTVKAGTNVCLDATV 12822
Qy 132 ----- 131
Db 12823 FGKPMPTVSWKKGDTLLKPAEGIKMAMORNLCLELFSVNRKDSGDTITAEENSSGSKSA 12882
Qy 132 ----- 131
Db 12883 TIKLVLDKPGPASVKINKMYSORAMLSWEPPLDGGSEITNYIVDKRETSRPNWAQVS 12942
Qy 132 ----- 131
Db 12943 ATVPITSGVEKLIBGHEYQFICAENKYGVDPVFTEPAIAKPNYDPPGRCDPPVISNI 13002
Qy 132 ----- 131
Db 13003 TKDHMTVSWKPPADGGSPITGYLLEKRETQAVNWKVNRKPIIERTLTKATGLQEGTEYE 13062
Qy 132 ----- 131
Db 13063 FRVTAINKAGPKPSDASKAAYARDPOYPAPPAPPKYVDITRSVSLSWGKPAYDGSP 13122
Qy 132 ----- 131
Db 13123 IIGYLVKRADSNWRCNLPNLQKTRFEVTGLMEDTQOFRVYAVNKIGYSDPSDVP 13182
Qy 132 ----- 131
Db 13183 DKHPKDIILIPPEGEHDADRLKTLILRAGVTMRLYVPVKGRPPKXTWSKPNVNLDRIG 13242
Qy 132 ----- 131
Db 13243 LDIKSTDFDIFRCENVNKYDAGKYILTLENSCGKKEYIVVVLDTGPPINVTVKEIS 13302
Qy 132 ----- 131
Db 13303 KDSAVVTWEPLIIDGSPPIINYVQKRAERKSWSTVTECSKTSFRVPNLEEGSKYFFR 13362
Qy 132 ----- 131
Db 13363 VFAENEYIGDPGETRDVAKASQTPGPVVDLKVRSVSKSCSIGWKKPHSGGSRIGYV 13422

Qy 132 -----NGKEY----- 136
Db 13423 VDFLTEENKQWQVMKSLSLQYSAKDLTEGKEYTFRSAENENEGCTPSEITVWARDVVA 13482
Qy 137 ----- 136
Db 13483 PDLDLKGLPDLCLYLAKENSFRKLIPIKGPAPSVSWKKGEDPLATDTRVSVESAVNTT 13542
Qy 137 ----- 136
Db 13543 LIVYDCQKSDAGKYITITLKNVAGTKEGTISIKVWGKPGIPTGPIKFEDEVTAEMTLKWP 13602
Qy 137 ----- 136
Db 13603 PKDDGSEITNYILEKRSVNKNKWTVCASAVQKTTFRVTRLHEGMEYTFRVSANKYGVG 13662
Qy 137 ----- 136
Db 13663 EGLKSEPIVARHPEDVPDAPPPPNIVDVHRDSVSLTWTDPKTKGGSPITGVHLEFKERNS 13722
Qy 137 ----- 136
Db 13723 LLWKRANKTPIRMDFKVTGLTEGLEYEFRVMAINLAGVGKPSLPSEPVVALDPIPPGK 13782
Qy 137 ----- 136
Db 13783 PEVINITRNSVTLIWTEPKYDGGHKLTCYIVEKRDLPSKSHMKANHVNVPECAFTVTDLV 13842
Qy 137 ----- 136
Db 13843 EGGKYEPRIRAKTAGAISAPSESTETIICKDEYEAPTVLDPTIKDGLTIKAGDTIVLN 13902
Qy 137 ----- 136
Db 13903 AISILGKPLPKSSWAKGDIRPSDITQITSTPTSSMLTIKYATRKDAGEYITITATNPF 13962
Qy 137 ----- 136
Db 13963 TKVEHVKTVDVPCPGPVEISNVAEKATLTWTPPLEDDGGSPKSYILEKRETSRLW 14022
Qy 137 ----- 136
Db 14023 TVSEDIQSRHVATKLIQONEYIFRVSANVHYKGEPVQSEPKMVMDFGPPGPKPE 14082
Qy 137 -----KCKVSN----- 142
Db 14083 VSNVTKNATVSWKRPVDDGGSEITGYHVERREKSLRWVRAIKTPVSDLCRCKVTGLOEG 14142
Qy 143 ----- 142
Db 14143 STYEFVSAENRAGIGPPSEASDSVLMKDAAYPPGPPSNPHVTDITTKKSASLANGKPHYD 14202
Qy 143 ----- 142
Db 14203 GGLEITGVVEHQVKGDEAWIKDITGTALRITQFVVPDLOTKEYNFRISAINDAGVGP 14262
Qy 143 -----KALPAP----- 148
Db 14263 AVIPDVEIVEREMAPDFELDAELRRTLVRAGLSIRIFVPIKGRPAPEVTWTKONINUKN 14322
Qy 149 ----- 148
Db 14323 RANIENTESFTLLIIPECNRYDTGKFVMTIENPAGKSGFVNVRLDTPGPVNLRLPTDI 14382
Qy 149 -----IEK----- 151
Db 14383 TKDSVTLHMDLPLIDGGSRTINYIVEKREATRKSYSTATTCKHCKTKYKVTGLSECEYFF 14442
Qy 152 ----- 151
Db 14443 RVMAENEYIGEPTEPVPKASEAPSPDLSINIMDITKSTVSLAWPKPHGSGSKITGY 14502

QY	152	-----	151
Db	14503	VIEAQRKSDOWHITFTVKGLECVVRNLTEGEYTFQVMVANSAGRSAPRESRPVKEQ	14562
QY	152	-----	151
Db	14563	TMLPELDLGGIYQKLVIAKAGONIKVEIPVLGRPKFTVTWKGDQILKQTORVNFETTAT	14622
QY	152	-----	151
Db	14623	STILNINECVSRSDGYPPLTARNIVGEVGDVITIQVHDIPGPGTGPVKFDEYSSDFVTF	14682
QY	152	-----	151
Db	14683	WDPPENDGGVPISNYVEMROTDSTTWELATTVIRTTYKATRLTTGLEYQFRVKAQRY	14742
QY	152	-----	151
Db	14743	GVGPGITSAMIYANYPFKVPGPGTPOVTAVTKDSMTISWHEPLSDGGSPILGYHVERKE	14802
QY	152	-----TISKA-----	156
Db	14803	RNGILMQTVSKALVPGNIFKSSGLTDGIAEYFRVIAENMAGKSPKSPSEPMALDPIDP	14862
QY	157	-----	156
Db	14863	PGKPVPLNITRHTVTLKWAKEPYTGFKITSYIVEKRDLPNGRWLKANFENEFVTS	14922
QY	157	-----	156
Db	14923	GLTEDAAYEYFRVIAKNAAGISPPSEPSDAITCRDDVEAPKIKVDVKDVTILKAGEAF	14982
QY	157	-----	156
Db	14983	RLEADYSGRPPPTMWSKDGKELEGTAKLEIKIADPSTNLVKNKDSRRDGGATLTATNP	15042
QY	157	-----	156
Db	15043	GGFAKHIFNVKVLDRPGPEGLAVTEVSEKCVLSWFPPLDDGAKIDHYIVOKRETSR	15102
QY	157	-----	156
Db	15103	LAMTNVASEVQVTKLVTKLLKNGEYIFRMAVNKYGVGELESEPVLAENVYGPDPKP	15162
QY	157	-----	156
Db	15163	NPEVTTITKDSMVVCHPDSGGSELIINYIVERRDKAGQWIKCNKKTILDLRYKVSGL	15222
QY	157	-----	156
Db	15223	TEGHEYFRIMAENAGISAPSPTSPFYKACDVFVKPGPGNPRVLDTSKSSISIAWKP	15282
QY	157	-----	156
Db	15283	IYDGSSEITGYMVEIALPEDEMOIVTPPAGLKATSYITIGLTENOQYKIRIYAMNSEGL	15342
QY	157	-----	156
Db	15343	GEPALVPGTPKAEDRMLPPEIELDADLRKVVTIRACCTLRFPVKGRDPEVKWARDHG	15402
QY	157	-----	156
Db	15403	ESLDKASIESASYTLIIVGNVNRFDGKYLTVENSNGSKSAFVNVRLDTPGPPQDLK	15462
QY	157	-----	156
Db	15463	VKEVTKTSVTLWDPPLLDGGSKIKNYIVEKRESTRKAYSTVATNCHKTSWKVDLQEGC	15522
QY	157	-----	156
Db	15523	SYIFRVAENEYIGILPAETAESVKASERPLPPGKITLMDVTRNSVLSWEKPEHDGGR	15582
QY	157	-----	156
Db	15583	ILGYIVEMQTKGSKWATCATVKVTEATITGLIQGEYSFRVSAQNEKGISDPQLSVPV	15642
QY	157	-----	156
Db	15643	IAKDLVIPPAFKLLFNFTFVLAGEDLKVDVPFICRPTPAVTHKDNVPLKQTRVNAEST	15702
QY	157	-----	156
Db	15703	ENNSLLTIKDACREDYGVVVKLTNSAGEAIELNVILDKPGPTGPVKMDEVADTSIT	15762
QY	157	-----	156
Db	15763	LSWGPVKYDGGSSINNYIVEKRDYTTTWTQIVSATVARTIKACRLKTCGEYQFRIAAEN	15822
QY	157	-----	156
Db	15823	RYGKSTYILNSEPTVAQYPPFKVPGPGTPVVTLSRDSMEVQWNEPISDGGSRVIGHLER	15882
QY	157	-----	156
Db	15883	KERNLSILWVKLNKTPIPTQTKFTTGLEEGVEYEFVRVSAENIVGIGKPSKVCYVARDPC	15942
QY	157	-----	156
Db	15943	DPGRPEALIVTRNSVTLQWKPTYDGGSKITGYIVEKKELPEGRWMKASFTNIIDTHFE	16002
QY	157	-----	156
Db	16003	VTGLVEDHRYEYFRVIAARNAAGVSEPSSESTGAIATARDEVDPPIRISMDPKYKDIVVHAGE	16062
QY	157	-----	156
Db	16063	SFKVDADIYKPIPTIOWIKGDOELSNATLEIKSTDFTATSLVKDAVRVDSNGYILKAK	16122
QY	157	-----	156
Db	16123	NVAGERSVTNVKVLDRPGPGPVVIGSVTAECTILAWKPLQDGGSDIINYIVERRET	16182
QY	157	-----	156
Db	16183	SRLVTVVDANVOTLSCKVTKLLEGNEYTPRIMAVNKNYGVGELESEPVVAKNPFVVPDA	16242
QY	157	-----	156
Db	16243	PKAPEVTTVTKDSMIVVWERPASDGGSEILGYVLEKRDKEGIRWTRCHKRLIGELRLVT	16302
QY	157	-----	156
Db	16303	GLIENHDYEFVRVSAENAAAGLSEPPSPSAYQKACDPIYKPGPPNNPKVIDITRSSVFLWS	16362
QY	157	-----	156
Db	16363	KPIYDGGCEIQGYIVEKCDVNVGEMTCTPPTGINKNTNIEVEKLEKHEYNFRICAINKA	16422
QY	157	-----	156
Db	16423	GVGEHADVPPIIVEKLEAPDIDLLELRKIINIRAGSLRLFPVPIKGRPTPEVKWGV	16482
QY	157	-----	156
Db	16483	DGEIRDAAIIDVTSFTSLVDNVNRYDSGKYTLTLENSGTSKSAFVTVRVLDTPTSPPVN	16542
QY	157	-----	156
Db	16543	LKVEITKDSVSTWEPPLDGGSKIKNYIVEKREATRKSAAVVTNCHNSMKIDQLQE	16602
QY	157	-----	156
Db	16603	GCSYIFRVAENEYIGLPAQTAADPIKVAEVPQPPGKITVDDVTRNSVLSWTKPEHDGG	16662
QY	157	-----	156

Db 16663 SKIIQIVEMQAHEKSEKARVKSQAVITNLITQGEYLFRRVAVNEKGRSDPRSLAV 16722
Qy 157 ----- 156
Db 16723 PIVAKDLVIEPDVKPAFSSYSVOVGDLKMEVPISGRPKPTITWTKDGLPLKQTRINTV 16782
Qy 157 ----- 156
Db 16783 DSLDLTTLISKETHKDDGGQGITVANVVGQKTASIEIWLTDKPPKGPVKFDDVSAES 16842
Qy 157 ----- 156
Db 16843 ITLSNPPLYTGGCQITNYIVOKRDTTITVNDVVSATVARTLKVTKLTGTEYQFRIFA 16902
Qy 157 ----- 156
Db 16903 ENRYGQSALESDDPIVAQYPPYKEPGPGTFFATAISKDSWVIQWHEPVNNGSPVIGHL 16962
Qy 157 ----- 156
Db 16963 ERKERNILWTKVNTIITHTQFAQNLEEGIEYFRVVAENIVGVGKASKNSECYVARD 17022
Qy 157 ----- 156
Db 17023 PCDPGTEPIMVKRNEITLQWTKPVYDGGSMITGYIVEKKRDLDPGRWMAKSFNTVETQ 17082
Qy 157 ----- 156
Db 17083 FTVSGLTEDQRYEFRVIAKNAAGAIKSPSDSTGTPITAKDEVELPRISMDPKFRDTIVNA 17142
Qy 157 ----- 156
Db 17143 GETFRLEADVHGKPLPTIEWLRGDKIEBESARCEIKNTDFKALLIVKDAIRIDGGQYILR 17202
Qy 157 ----- 156
Db 17203 ASNVAGSKSPVNVKVLDRPPEGPVQVTGVTSEKSLTWSPPLODGGSDISHYVVEKR 17262
Qy 157 ----- 156
Db 17263 ETSRLAWTVASEVNTNSLVTKLEGNEYVFRIMAVNKYGVGEPLASAPVLMKNPFVLP 17322
Qy 157 ----- 156
Db 17323 GPPKSLEVTNTAKDSMTVCWNRPSDGGSEIIGYIVEKKRDRSGIRWIKCNKRRTDLRLR 17382
Qy 157 ----- 156
Db 17383 VTGLTEDHEYFRVSAENAGVGEPSATVYIYKACDPVKPGPPTNAHIIVDTTKNSITLA 17442
Qy 157 ----- 156
Db 17443 WCKPIYDGGSEILGYVVEICKADEEWQIVTPQTGLRVTREISKLTEHQEKIRVCALN 17502
Qy 157 ----- 156
Db 17503 KVLGEATSVPGTVKPEDKLEAPELDSELKRGIVRAGGSARIHIHPFKGRPMPEITWS 17562
Qy 157 ----- 156
Db 17563 REEGFTDKVQIEKGVNTQLSIDNCDNRDAGKYILKLENSGSKSAFVTVKVLDTGPGP 17622
Qy 157 ----- 156
Db 17623 QNLAVKEVKDSAFVWEPPIIDGKAVKNYVIDKRESTRKAYANVSSKCSKTSFKVENL 17682
Qy 157 ----- 156
Db 17683 TEGAIYFRVMAENBFGVGPVETVDVAKAAEPPSPGKVLTDVQSOTASLWKEKPEHD 17742
Qy 157 ----- 156
Db 17743 GGSRLVGVVEMQPKGTEKWSIVAESKVCNAVVTGLSSGQEQYQFRVAKYNEKKGSDPRVL 17802

Qy 157 ----- 156
Db 17803 GVPVIAKDLTITQPSLKLFPNTYISIOAGEDLKIEIPVIGRPRNISWVKDGEPLKQTRVN 17862
Qy 157 ----- 156
Db 17863 VEETATSTVLHIKEGNKDDFGKYTVTATNSAGTATENLSVIVLEKPGVPVGRFDEVSA 17922
Qy 157 ----- 156
Db 17923 DFVVISWEPAYTGGCQISNYIVEKRDTTTHHMSATVARTTIKTKLTGTEYQFRI 17982
Qy 157 ----- 156
Db 17983 FAENRYGKSAPLDSKAVIVQYPPFEKPGPGTFFVTSISKDQMLVQWHEPVDGGTKIIGY 18042
Qy 157 ----- 156
Db 18043 HLEQKEKNSILWVKLNKTPIQDTKFKTTGLDEGLEVEFKSAENIVGIGKPSKVECFVA 18102
Qy 157 ----- 156
Db 18103 RDPDPPGRPEAIVITRNNVTLMKKKPAYDGGSKITGYIVEKKDLDPGRWMAKSFNTVLE 18162
Qy 157 ----- 156
Db 18163 TEFTVSGLVEDQRYEFRVIAARNAAGNFSEPSDSSGAIITARDEIDAPNASLDPKYKDVIVV 18222
Qy 157 ----- 156
Db 18223 HAGETFVLEADIRKPIPDVVWSKDGKELEETAARMEIKSTQKTTLVVKDCIRTDGGQY 18282
Qy 157 ----- 156
Db 18283 ILKLSNVGGTSPITVYKVLDRPSPGEPGLKVTGVTAEKCYLANNPLODGGANISHYII 18342
Qy 157 ----- 156
Db 18343 EKRETSRLSWTQVSTEVQALNYKVTLLPGNEYIFRVMVAVNKYGIGEPLESGPVTACNPY 18402
Qy 157 ----- 156
Db 18403 KPPGPPSTPEVSATKDSMVVTWARPVDGGTEIEGYILEKRDKEGVRWTKCNKTLTDL 18462
Qy 157 ----- 166
Db 18463 RLVRTGLTEGHSYEFERVAENAGVGEPSVIFYRACDALYPPGPPSPNPKVTDTSRSSV 18522
Qy 167 ----- 166
Db 18523 SLAMSKPIYDGGAPVKGYVVEVKEAADEWTTCTPPTGLOGKQFTVTKLKENTENFRIC 18582
Qy 167 ----- 169
Db 18583 AINSEGVGEATLPGSVVAQERIEBPPEIELDALRKVVVLRASATLRLFVTIKGRPEPEV 18642
Qy 170 ----- 169
Db 18643 KWKAEGILTDRAQIEVTSSTMLVIDNVTRFDSGRYNLTLENNSGSKTAFVNVRLDSP 18702
Qy 170 ----- 169
Db 18703 SAPVNLATIREVKDSVTLSEWPEPLIDGKAKITNYIVEKRETTKRAYATITNNCTKTTFRI 18762
Qy 170 ----- 169
Db 18763 ENLOEGCSYFVRVLASNEYIGLPAETTEPVVKVSEPPPLPGRVTLVDVTRNTATIKWEKP 18822
Qy 170 ----- 169
Db 18823 ESDGSKITGYVVEVEMQKGEKSWSTCQVKTLEATISGLTAGEEYVFRVAAVNEKGRSDP 18882

QY	170	-----	169
Db	18883	RLQGVPIARDIEIKPSVELPPHTENVKAREQLKIDVPPKGRQATVNNRKDGQTLKEIT	18942
QY	170	-----	169
Db	18943	RNVSSKVTSLSIKEASKEDVGYELCVNSAGSITVPITIIIVLDRPFGPPGPIRIDEV	19002
QY	170	-----	169
Db	19003	SCDSIITISWNPPEYDGGCQISNIVVEKKETTSTTHIVSQAVARTSIKIVRLTTGSEYQF	19062
QY	170	-----	169
Db	19063	RVCAENRYKSSYSESSAVVAEYFPFPGPGTPKVHHATKSTMLVTWQVPVNDGSRVI	19122
QY	170	-----	169
Db	19123	GYHLEYKERSIILWSKANKILIAADTOVKVSGLDEGLMEYRYVAENIAGIGKSKCEPV	19182
QY	170	PSRD-----	173
Db	19183	PARDPCDPGQPEVTNITRKSVLSKSKPHYDGGAKITGYIVERRELDPGRWLKCNVTNI	19242
QY	174	-----	173
Db	19243	OETYFEVTELTEDQRYEYFVAFARNADSVSESESTGPIIVKDDVEPPVMMOVKFRDVI	19302
QY	174	-----	173
Db	19303	VKAGEVLKINADIAGRPLPVISWAKDGEIERARTEIISTDNHTLLTVKDCIRRDGTQ	19362
QY	174	-----	173
Db	19363	YVLTAKNAVAGTRSAVAVCKVLDKPGPPAGPLEINGLTAEKCSLSWGRPQDGGADIDYH	19422
QY	174	-----	173
Db	19423	RKKRETSHLAWTICEGELQWTSCKVTLLKGNEYIFRVTGVNKYGVGEPLSVAIKALDP	19482
QY	174	-----	173
Db	19483	FTVPSPPTSLEITSVKESMTLWSPESDGGSEISGYIIERREKNSLRVRVNNKRPVYD	19542
QY	174	-----	173
Db	19543	LRVKSTGLREGCEYEVVAENAGLSLPSLSETPLIRAEDPVFLPSPPKIVDSGKTT	19602
QY	174	-----	173
Db	19603	ITIAWVKPLFDGGAPITGYTVVEYKKSDTDWKTISIQLRGTEYTIISGLTGAEYVFRVKS	19662
QY	174	-----	173
Db	19663	VNVKASDPSDSDPOIAKEREEREEPLFDIDSEMRKTLIVKAGASFMTVPFGRGPVNL	19722
QY	174	-----	173
Db	19723	WSPKPTDLRTRAYVDTDSRTSLTIENANENDSGKYTLTIONVLSAASLTLVVKVLDTPG	19782
QY	174	-----	173
Db	19783	PPTNITVQDVTKESAVLSMDVPENDGGAPVKVNHIEKREASKKAWSVTNNCNRLSYKVT	19842
QY	174	-----	173
Db	19843	NLOEGAIYYFRVSGENEFVGIPAEITKEGVKITEKPSPEKLGVTISISKDSVSLTWLKE	19902
QY	174	-----	173
Db	19903	HDGGSRLVHVVALEKGQKNWYKCAVAKSTHHVSVGLSENSEYFFRVFAENQAGLSDPR	19962
QY	174	-----	173

Db	19963	ELLLPVLKEQLEPPEIDMKNFPSHTVYVRAGSNLKVDDIPISGKPLPKVTLSRDGVPLKA	20022
QY	174	-----	173
Db	20023	TMRFNTEITAENLTINLKESVTADAGRYEITAANSSTTTAFINIVVLDPRGPTGPVI	20082
QY	174	-----	173
Db	20083	SDITEBSVTLKWPPKYDGGSOVTNYILLKRETSTAVTTEVSATVARTMMKVMKLTGEE	20142
QY	174	-----	173
Db	20143	YQFRIKAENRFGISHIDSACVTVKLPYTPPGPSTPWVNTRESITVGWHEPVSNNGS	20202
QY	174	-----	173
Db	20203	AVGYHLEMKDRNSILWQANKLVIRTHFKVTTISAGLIYEFVRVYAENAGVKPSPHS	20262
QY	174	-----ELTKNOVSL-----	182
Db	20263	EPVLAIDACEPPRNVRTIDISKNSVLSWQOPAFDGGSKITGYIVERRDLPDGRWTKASF	20322
QY	183	-----TCL-----	185
Db	20323	TNVTETQFTISGLTQNSQYEFVRVAFARNAGVSGISNPSSEVVGPIITCIDSYGGPVIDPLEYT	20382
QY	186	-----	185
Db	20383	EVVKYRAGTSVKLRAGISGKPAPPTIEWYKDDKELQTNALVCVENTDILASILIKDADRLN	20442
QY	186	-----	185
Db	20443	SGCYELKLNRNMAASATIRVQILDKPGPPGPIEFKVTAEKITLLWRPPADGGAKIT	20502
QY	186	-----	185
Db	20503	HYIVEKRETSRVVSMVSEHLEECIITTKIKGNEYIFRVRVAVNKYGIGEPLESDSVA	20562
QY	186	-----	185
Db	20563	KNAPVTPGPGPIPEVTKITKNSMTVVWSRPIADGGSDISGYFLEKRDKKSGLGWFKVLKET	20622
QY	186	-----	185
Db	20623	IRDRQKVTLGTENSQYRVCVAVNAAGQFPSEPFSEFYKAADPIDPPGPAKIRIADST	20682
QY	186	-----	185
Db	20683	KSSITLWSKPVDGGSVAVTGYVVEIRQGBEEWTTVSTKGEVTTTEYVYVSNLKPGVNY	20742
QY	186	-----	185
Db	20743	FRVSAVNCAGQGEPIEMNEPVQAKDILEAPEIDLVAALRTSVIAKAGEDVQVLIPFKGRP	20802
QY	186	-----	185
Db	20803	PPTVTRKDEKLGSDARYSIENTDSSLLTIPQVTRNDTGKYLITTIENGVGEPKSTVS	20862
QY	186	-----	185
Db	20863	VKVLDTAAQCKLVKHVSRGTVTLWDPPLIDGGSPIINVIEKRDATKRTKTSVSHKC	20922
QY	186	-----	185
Db	20923	SSTSFKLIDLSEKTPFFRVLAEINEIGICEPCTETPEVKAEEVPAPIRDLMSKDSTKTSV	20982
QY	186	-----	185
Db	20983	ILSWTRPDFDGGSVITEYVVERKKGEBQWTHSHAGISKTCIEVYSOLKEQSVLEFRVFAKN	21042
QY	186	-----	185

Db 21043 EKGLSDPVTIGPITVKELIITPEVDLSDIPGAQVTVRIGHNVHLELPYKGPSPISWLK 21102
Qy 186 ----- 185
Db 21103 DGLPLKESEFFVRSKTEKNITILSIKNAKEHGGKYTVILDNAVCRVIAPIVITILGPSPK 21162
Qy 186 ----- 185
Db 21163 PKGPIRFDEIKADSVILSWDPEDNGGGEITCYSIEKRETSQTNKMKVCSSVARTFKVP 21222
Qy 186 ----- 185
Db 21223 NLVKDAEQFRVRAENRYGVSOPLVSSIIVAKHOFRIIPGPGKXPVIYNWTSOGMSLTWDA 21282
Qy 186 -----VKGFY----- 190
| | |
Db 21283 PVDGSEVTGFHVEKKERNILMQKVNTSPISGREYRATGLVEGLDYQFRVYAENSAGL 21342
Qy 191 ---PSD----- 193
| | |
Db 21343 SSPSPDKFTLAVSPDPGTPDYIDVTRETITLKNPPLRDGGSKIYGSIEKROGNER 21402
Qy 194 ----- 193
Db 21403 WVRCNFTDVSECQYTVTGLSPGDRYEFRIIARNAVGTISPPSQSGIIMTRDENVPPIVE 21462
Qy 194 ----- 193
Db 21463 FGPEYFDGLLIKSGBSRLIKALVQGRPVPRVTWFKDGVGEIEKRMNMEITNVLGTSFLVR 21522
Qy 194 ----- 193
Db 21523 DATDRHGVYTVAKNASGSAKAEIKVKVQDTPGKVGPIRFTNITGKMTLWMDAPLND 21582
Qy 194 ----- 193
Db 21583 GCAPITHYIEKRETSRLAWALIEDKCAQSYTAIKLINGNEYQFRVSAVNFVGRPLD 21642
Qy 194 ----- 193
Db 21643 SDPVVAQIQYTVDPAPGIPESNITGNSITLTWARPESDGGSEIQOYLIERREKKSTRWV 21702
Qy 194 ----- 193
Db 21703 KVISKRPISETRFRKVTGLTEGNEYEFHVMAENAGVPGSGISRLIKREPVNPPGPTV 21762
Qy 194 -----TAVW----- 198
: : :
Db 21763 .VKVDTSTVTSLEWSKPVFDGMEIIGYIEMCKTDLDGMHKVNAEACVKTXYTVDLQ 21822
Qy 199 ----- 198
Db 21823 AGEYKFRVSAINGAGKDSCEVTGTIKAVDRLTAPELDIDANFKQTHVVRAGASIRLFI 21882
Qy 199 -----ESN----- 201
| | |
Db 21883 AYQGRPTTAVMSKPDNSLRLADIHTDTSFTLTVCNCRNDAGKYTLTVENNNGSKSI 21942
Qy 202 ----- 201
Db 21943 TFTVKVLTPGPGPITPKDVTGRSATLMDAPLLDGGARITHHYVVEKREASRRSQVIS 22002
Qy 202 ----- 201
Db 22003 EKCTRQIFKVNDLAEGVPYIFRVSANVEYGYGEPEYEMPEPIVATEQPAPPRLDVVDTSK 22062
Qy 202 ----- 201
Db 22063 SSAVLAWLKPDDGSRITGYLLEMROKQSDLWVEAGHTKQLTFTVERLVEKTEYEFVAK 22122
Qy 202 ----- 201
Db 22123 AKNDAGYSEPREAFSSVIIKEPQIEPTADLTGINTQLITCKAGSPFTIDVPIGRPAKV 22182

Qy 202 ----- 201
Db 22183 TWKLEMLRKETDRVSITTTKDRUTTLTVKDSMRGDSGRYFLTLENTAGVKTFSTVTVVIG 22242
Qy 202 -----QOP----- 204
| | |
Db 22243 RGPVTPGPIEVSSVSAESCVLWGPEDKGGGTETITNYIVEKRESGTTAQLVNSSVKRTQ 22302
Qy 205 ----- 204
Db 22303 IKVTHLTXYMEYSFRVSSSENFVSKPLESAPIIAEHFPVPSPAPTREPVHVHSANMSI 22362
Qy 205 -----ENNYKT----- 211
| | |
Db 22363 RWEOPYHDGSKIIGYWVEKKERNITLWVKENKVPCLCNKYKTVGLVEGLEYOFTYALN 22422
Qy 212 ----- 211
Db 22423 AAGVSKASEASRPIMAQNPDAPGRPEVTDVTRSTVSLWSAPAYDGGSKVVGYIIERKP 22482
Qy 212 ----- 211
Db 22483 VSEVGDGRWLKCNKNTIIVSDNFTVTALSEGDTYEFRLAKNAAGVISKSESTGPTCRD 22542
Qy 212 ---pp----- 213
| |
Db 22543 EYAPKAEELDARLHGLDVTIRAGSDLVLDAAVGGKPEPKIITWTKGDKELDKCVSLQYT 22602
Qy 214 -----VLDSDG----- 219
| | |
Db 22603 GKRAVIFKCDRSDSGRYTLTVKNASGTAKAVSMVKVLDSPGCGKLTVSRVTOEKCTL 22662
Qy 220 ----- 219
Db 22663 AWSLPQEDGGAETHYIVERRETSRLNWIVVEGECPTLSYVVYTRLIKKNIEYFRVAVNK 22722
Qy 220 -----SFFL----- 223
| | |
Db 22723 YGPGVPESEPIVARNSTIPSPGCIPEEVGTGKEHIIQWTKPESDGNESINYLVOKR 22782
Qy 224 ----- 223
Db 22783 EKESLRWTRVNKYVYVDTRLKVTSLMEGCDYQFRVTAANAENSEPSESNFISCREPS 22842
Qy 224 ----- 223
Db 22843 YTPGPPSAPRVVDTTKHSISLAWTKPMYDGGTDIVGYVLEMQEKDQDQWYRVHTNATRN 22902
Qy 224 ----- 223
Db 22903 TEFTVPDLKMGQYSFRVAANVKGMSEYSESIAEIEPVERIEIPDLELADDLKKTVTIR 22962
Qy 224 -----YSKLTVDK----- 231
| | |
Db 22963 AGASLRMLVSVGRPPPVITWKSOGIDLASRAIIDTTESYSLLIYVDKYNRYDAGKYTIEA 23022
Qy 232 ----- 231
Db 23023 ENQSGKKSATVLVKVYDTPGPCPSVKVEVSRDSVTITWEIPTIDGGAPINNYIVEKREA 23082
Qy 232 ----- 231
Db 23083 AMRAFKTVTTKCKTLYRISGLVEGTMHYFRVLPENIYIGIGPCETSDAIVLSEVPLVPA 23142
Qy 232 ----- 231
Db 23143 KLEVVDVTKSTVTLAWEKPLYDGGSRLTGYVLEACAGTERMMKVTLKPTVLEHTVSL 23202
Qy 232 ----- 231
Db 23203 NEGEQYLFRIRAQNEKGVSEPRETVTAVTVQDLRVLPTIDLTMPQKTIHVPAGRPVELV 23262

QY 232 ----- 231
Db 23263 IPIAGRPAAAWFAGSKLRESERVVETHIKVAKLTIRETTIRDTGEYTLLELKNVTGT 23322
QY 232 ----- 231
Db 23323 TSETIKVIIIDKPGPTGIKIDEIDATISITISWEPPELGGAPLSGYVVEQDAHPGW 23382
QY 232 ----- 231
Db 23383 LPVSESVRTSTFKFTRLTEGNEVFRVAATNRFPGISYLOSEVIECRSIRIPGPPETLQ 23442
QY 232 ----- 231
Db 23443 IFDVSRDGWTLTWYPPEDDGSQVTGYIVERKEVRADRVVRKNKVPVTWTRYRSTGLTSG 23502
QY 232 ----- 231
Db 23503 LEYHRVTAINARGSGKPSRSPKPIVAMPDIAPPGKPNPRVTDTRTSVSLAWSPEDE 23562
QY 232 ----- 231
Db 23563 GSKVTGYLIEQKVDQHEWTKNTTPTKIREYTLTHLPQGAERYRVLACNAGGPGEPA 23622
QY 232 ----- SRWQQ- 237
Db 23623 EVPGTVKVTEMLEYDPYELDERYQEGIFVRQGVIRLTPIKGPFPICKWTKEGODISK 23682
QY 238 ----- 237
Db 23683 RAMIATSETHELVIKEADRGSGTYDLVLENKCGKAVYIKVRVIGSPNSPEPLEYDD 23742
QY 238 ----- 237
Db 23743 IQVRVSVWRPPADGGADILGYILERREVPKAAWTTIDSRVGTSLVVGLKENVEYH 23802
QY 238 ----- 237
Db 23803 FRVSAENQFISGPLKSEBPVTKPLNPPPEPPNPPELDVTKSSVLSWSRKPDDGGS 23862
QY 238 ----- 237
Db 23863 RVTGYIERKETSTDKVVRHNKTQITTTWYTVGLVPDAEYFRIIAQNVDVLSETSPAS 23922
QY 238 ----- 237
Db 23923 EPVCKDPDKPSQGELEILSISKDSVTLOWEKPCCDGGKEILGYWVEYRQSGDSAWKK 23982
QY 238 ----- 237
Db 23983 SNKERIKDKQFTIGGLLEATEYEFVFAENETGLSRPRRTAMSIKTLTSGEAPGIRKEM 24042
QY 238 ----- 237
Db 24043 KDVTTLGEAALSCQIVGRPLPDIKWYRFGKELIQSRKYMSSDGRTHTLVTMTTEEQED 24102
QY 238 ----- 237
Db 24103 EGYVTCIATNEVEVETSSKLLLOATPQFHPGYPLEKEYYGAVGSLRLHVMYIGRPVPA 24162
QY 238 ----- NVF- 240
Db 24163 MTWFHGOKLQNSENTIENTEYTHLVMKNVQRKTHAGKYKVLNSVFGTVDAILDVEI 24222
QY 241 ----- 240
Db 24223 QDKPKDPTGPIVIEALLKNSAVISWKPADGGSWITNYYVEKCEAKEGAQWLVSSAIS 24282
QY 241 ----- 240
Db 24283 VITCRIVNLNENAGYFRVSAONTFGISDPLEVSSVVIKSPFKPGPKPTITAVTKD 24342
QY 241 SCSV ----- 244

Db 24343 SCVWAMPPASDGGAKIRNYILEKREKKONKWISVTTEEIRETVFSVKNLIEGLEVEFRV 24402
QY 245 ----- 244
Db 24403 KCENLGSESEWSEISEPITPKSDVPIQAPHFKEELNLRVYQSNATLVCVKVTHGPKPIV 24462
QY 245 ----- 244
Db 24463 KWYRQKEIIADGLKYRIOEFKGGYHQLIIASVTDDATVYQVRATNQGSGVSSTASLEV 24522
QY 245 ----- 244
Db 24523 EYPAKIHLPKLTLEGMGAHVHALRGEVVSIKIPFSGKPDVITWQKGODLIDNNGHYQVIVI 24582
QY 245 ----- 244
Db 24583 RSFTSLVPFNGVERKDAGFYVVCANKRFGIDQKTVELDVADVPDPGRGVKVSASRDSVN 24642
QY 245 ----- 244
Db 24643 LIWTEPASDGGSKITNYIVEKCATTAERWLRVGOARETRYTVINLFGKTSYQFRVIAENK 24702
QY 245 ----- 244
Db 24703 FGLSKPSEPEPITTKEDKTRAMNYDEVDREVSMTKASHSTKELYEKYMIABDLGR 24762
QY 245 ----- MHEA- 248
Db 24763 GEGFIVHRCVETSSKTYMAKFVKVKGTDQVLVKEISILNIAHRNHLHLESFESMEE 24822
QY 249 ----- LHNH- 252
Db 24823 LYMWIFFISGLDIFERINTSAFELNEREIVSYVHVCEALQFLHSHNIGHFDIRPENIY 24882
QY 253 ----- 252
Db 24883 QTRRSSTIKIIEFGARQLKPGDNFRLFTAPEYAYEYVHQHDVSTATDMNSLGLTVVY 24942
QY 253 ----- YT- 254
Db 24943 LLSGINPFLAETNQIENIMNAEYTFDEAFKEISIEAMDVFDRLLVVERKSRMTASEA 25002
QY 255 ----- QKSLSL- 261
Db 25003 LQHPWLKQKIERVSTKVIRTLKRRYYHTLIKKDLNMYVSAARISCGGAIRSOKGYSVAK 25062
QY 262 ----- 261
Db 25063 VKVASTIEIGPVSQIMHAVGEGGHVKYVCKIENYDQSTQVTWYFGVROLENSEKEYEITY 25122
QY 262 ----- 261
Db 25123 EDGVAILYVKDITKLDGTYRCKVNVNDYGEDSSYABELFKGVREYDYVYCRRTMKIKRR 25182
QY 262 ----- 261
Db 25183 TDTMRLLERPPEPTLPYLNKTAAYGVENVRFGVTITVHPPEVHTWYKSGQIKPGDNKKY 25242
QY 262 ----- 261
Db 25243 TFSKGLYQLTINSVTDDAEYTVVARNKYGEDSCAKKLTVLTHPPPTDSTLRPMFKR 25302
QY 262 ----- 261
Db 25303 LLANAEQEQSQVCFEIRVSGIPPTLKWEKDGQPLSLGNPIEIIHGLDYALHIRTDL 25362
QY 262 ----- 261
Db 25363 PEDTGYRYVTATNTAGTSCQAHQLQVERLRYKKQEFKSEERHVKQIDKTLRMAEIL 25422
QY 262 ----- 261

Db 25423 SCTESVPLTQVAKREALREAVLYKPAVSTKTVKGFRLEIEEKKKEERKLMPDYDVPEPRK 25482

QY 262 -----PGK 264
|||

Db 25483 YKOTTIEEDQRIKQVPMNDKMKWKKIRQYEMPGK 25518
|||

RESULT 15

Q8WZB3
ID Q8WZB3 PRELIMINARY; PRT; 26926 AA.
AC Q8WZB3;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE N2B-titin isoform.
GN TTN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20309627; PubMed=10850961;
RA Freiburg A., Trombitas K., Hell W., Cazorla O., Fougereusse F.,
RA Centner T., Kolmerer B., Witt C., Beckmann J.S., Gregorio C.C.,
RA Granzier H., Labeit S.;
RT "Series of exon-skipping events in the elastic spring region of titin
RT as the structural basis for myofibrillar elastic diversity.";
RL Circ. Res. 86:1114-1121(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21573839; PubMed=11717165;
RA Bang M.L., Centner T., Fornoff F., Geach A.J., Gotthardt M.,
RA McNabb M., Witt C.C., Labeit D., Gregorio C.C., Granzier H.,
RA Labeit S.;
RT "The complete gene sequence of titin, expression of an unusual ~700
RT kDa titin isoform and its interaction with obscurin identify a novel
RT Z-line to I-band linking system.";
RL Circ. Res. 89:1065-1072(2001).
DR EMBL; AJ277892; CAD12455.1;
DR InterPro; IPR000282; Cytochrome receptor_2.
DR InterPro; IPR000719; Euk.pkinase.
DR InterPro; IPR000577; FGGY_kin.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR002016; Peroxidase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00041; fn3; 132.
DR Pfam; PF00047; ig; 91.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00060; FN3; 132.
DR SMART; SM00409; IG; 111.
DR SMART; SM00408; IGG2; 93.
DR SMART; SM00410; IG_like; 3.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TykC; 1.
DR PROSITE; PS00933; FGGY_KINASES_1; UNKNOWN_1.
DR PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
SQ SEQUENCE 26926 AA; 2993285 MW; 169AB42637A7C1FB CRC64;

Query Match 33.9%; Score 485; DB 4; Length 26926;
Best Local Similarity 0.7%; Pred. No. 0.0002;
Matches 177; Conservative 44; Mismatches 42; Indels 23713; Gaps 55;

QY 2 WLK----- 4
|||

Db 1543 WLKNSDIIVPHKPKIRIEGKTGEAALKIDSTVSQDSAWYTATINKAGRTDTRCKVNV 1602

QY 5 -----AFYDK----- 9
|||

Db 1603 VEFAPPEPERKLIIPRGTYRAKEIAAPELEPLHLRYGOEWEGDLYDKERQKQKPPFKK 1662

QY 10 ----- 9

Db 1663 LLSLRKRGFGPAHFECRLTPIGDPTMVVWHLHDGKPLEAANRLRMINEFGYCSLDYGVA 1722

QY 10 ----- 9

Db 1723 SRDSGIITCRATNKVGTDTHTSATLIVKDEKSILVESQLPEGRKGILQRIEELERMAHEGAL 1782

QY 10 ----- 9

Db 1783 TGVTTDOKEKOKPDIVLYPEPVRVLEGETARFCRCVTGTPQPKVNWYLNGLIRKSRFR 1842

QY 10 ----- 9

Db 1843 VRYDGIHYLDIVDCKSYDTEGEVKVTAENPEGVIEHKVKLEIQOQREDFRSLRRAPPEPR 1902

QY 10 ----- 9

Db 1903 FHVHEPGKLQFEVOKVDRPVDTTETKEVVKLKRAERTIHEKVPSESELRSKFKRRTBEG 1962

QY 10 ----- 9

Db 1963 YYEAITAVELKSRKDESYELLRRKTKDELHLHWTKEELKALAEKGKITITFTPKDK 2022

QY 10 ----- 9

Db 2023 IELSPSEAPKIFERIOSQTVGQSDAHFRVVRVVGKPDPECEWYKNGVKIERSRIYVW 2082

QY 10 ----- 9

Db 2083 PEDNVCGLVIRDTAEDSASIMVKAINIAGETSSHAFLVQAKLIITFQELQDVVAK 2142

QY 10 ----- 9

Db 2143 DTMATFECETSEPPFVKVYKDGMEVHEGDKYRHSORKVHFLSILTIDTSDAEDYSCVL 2202

QY 10 ----- 9

Db 2203 VEDENVTTAKLIVGAVVEFVKELQDIEVPEYSYSGELEICIVSPENIEGKWHNDVELKS 2262

QY 10 ----- 9

Db 2263 NGYTIITRRGRQNLTVKDVTKEDQGEYSFVIDGKTKTKLKKMPRIAILQGLSDQKVC 2322

QY 10 ----- 9

Db 2323 EGDIVQLEVKVLSVESVEGVMKQGEVQPSDRHVHIVDKOSHMLLIEDMTKEDAGNYSFT 2382

QY 10 ----- 9

Db 2383 IPALGLSTGRSVSVYVDVITPLKDVNVIBGTRKAVLECKVSPDVTSVKWLNDQIKPD 2442

QY 10 ----- 9

Db 2443 DRVQAIKGTQKRLVINRTHASDEGPYKLIIVGRVETNCNLSVEKIKIRGLRDLTCTETQ 2502

QY 10 ----- 9

Db 2503 NVFVEVELSHGIDVLNFKDKKEIKPSSKYKIEAHGKIYKLTVLNMMKDDGKTYTFYAGE 2562

QY 10 ----- 9

Db 2563 NITSGKLTAVAGGAIKPLTDQTVAESQEAQVFECEVANPDSKGWLRDGHPLPLTNIRSE 2622

QY 10 ----- 9
Db 2623 SDGHRRLIATAKLDIGETTYKATSKTSAKLKVEAVKIKKTLKNTLVTEQDAVTV 2682
QY 10 ----- 9
Db 2683 ELTHPNVGVQWIKNGVWLESNEKYAISVKGTTIYSLRIKNAICVDESIVGFRGLGASA 2742
QY 10 ----- 9
Db 2743 RLHVEIVKIIKKPKDVTALENATVAFEVSVDHTVPVKHFKHVEIKPSDKHRLSERKV 2802
QY 10 ----- 9
Db 2803 HKMLQNIQSPDAGEYAVVQLECKAKLFEVTLHITKTMKNIETPETKTATSECEVSHF 2862
QY 10 ----- 9
Db 2863 NVPSMWLKNQVEIEMSEKFIYVQGLHQLIIMNTSTEDSAEYTFVCGNDQVSATLTVTP 2922
QY 10 ----- 9
Db 2923 IMITSMKLDINAEKDTITFEVTNVYEGISYKWLKNGVEIKSTDKCOMRTKKLTHSLNIR 2982
QY 10 ----- 9
Db 2983 NVHFGDAADYTFVAGKATSTATLYVEARHTEFRKHIDIKVLEKKRAMFECEVSEPDITV 3042
QY 10 ----- 9
Db 3043 QNMKDDOQLOITRIKIQKEKYVHRLIIPSTRMSDAGKYTVVAGNVSTAKLFEVGRDVR 3102
QY 10 ----- 9
Db 3103 IRSIKKEVQVIEQORAVVEFEVNEDDVAHWYKDGIENFQVQERHKYVVERRIHRMFIS 3162
QY 10 ----- 9
Db 3163 ETRQSDAGEYTFVAGNRSSVTLVNAPEPPVQLQELPVTYVQSGKPARCAVISGRPOP 3222
QY 10 ----- 9
Db 3223 KISWYKEQLLSTGFKFLHDGOEYTLLEIAPPDAAYTCEAKNDYGVATTSASLSV 3282
QY 10 ----- 9
Db 3283 EVPEVSPQOEMPVYPPAIITPLQDVTSEGQPARFCRVSGTDLKVSWSKDKKIKPSR 3342
QY 10 ----- 9
Db 3343 FFRMTQFEDTYQLEIAEAYPEDEGTYTFVNASNAVQVSSSTANLSLEAPESILHERIEQEI 3402
QY 10 ----- 9
Db 3403 EMEMKBFSSFLSAEBGLHSAELQSKINETLELLESVPYSTKFDSEKGTGPIFIKE 3462
QY 10 ----- 9
Db 3463 VSNADISMGDVATLSVTVIGIPKPKIQWFNGVLLTPSADYKFVFDGDDHSLIILFTKLE 3522
QY 10 ----- 9
Db 3523 DEGETCMASNDYGKTCISAYLAKINSKGEGHKDTESAVAKSLEKLGCPHFLKELK 3582
QY 10 ----- 9
Db 3583 PIRCAOGLPAIFEYTVVGEPAFTVTFWFKENKQLCTSVYITIHNPNSGGTFIVNDPQRED 3642
QY 10 ----- 9
Db 3643 SGLYICKAENMLGESTCAAEPLLLEDTMDTDPCKAKSTPEAPDPOTPLKGAPEAL 3702
QY 10 ----- 9

Db 3703 DSEQIATFVKDTILKAALITEENQOQSYEHIAKANELSSQLPLGAQELQSILOQDKLTP 3762
QY 10 ----- 9
Db 3763 ESTREFLCINGSIHFOPLKPESPNLOQIVSQKTSFSGKILMPPEPETOAVLSDTEKIF 3822
QY 10 ----- 12
Db 3823 PSAMSTEQINSLTVEPLKTLAEPEGNYPOSSIEPPMHSYLTVAEVLSPKEKTVSDTN 3882
QY 13 ----- 12
Db 3883 REQRTVLQGEAQASALILSQSLAEGHVESLQSPVMISQVNYBPLVPSEHSCTEGGKILI 3942
QY 13 ----- 12
Db 3943 ESANPLENAGQDSAVRIEKGSLRFFPLALEEKQVLLKEHSDNVVMPDQIIESKREPVA 4002
QY 13 ----- 12
Db 4003 IKKQVQVQGRDILLSKESLISGIPPEQRLNLKIQICRALQAAVASEQPGLESEWLNIEKV 4062
QY 13 ----- 12
Db 4063 EVEAVNITQEPRHIMCYLVTSAKSVTEETIIIEDVDPMQANLKMELRDALCAIIVEEI 4122
QY 13 ----- 12
Db 4123 DILTAEPRIQOGAKTSLQEMDSFSGSQKVEPITEPEVESKYLIISPEEVSYNFVQSRVK 4182
QY 13 ----- 17
Db 4183 YLDATPVTGCVASAVVDEKQDESLEKPESEKESSESSESTEEVATVKIQEAGSGFIKEDG 4242
QY 18 ----- 17
Db 4243 PMIHTPLVDTVSEBGDIVHLTTSITNAKEVNWYFENKLVPSDEKFKCLQDQNTVTLVDK 4302
QY 18 ----- 17
Db 4303 VNTEHQGEYVCEALNDSGKTATSAKLTVVYKRAAPVIKRKIEPLEVALGHLAKFTCEIQS 4362
QY 18 ----- 23
Db 4363 APNVRQWFKAGRIEYESDKSIRSSKYISSLEILLRTQVVDCGEYTCKASNEYGSVCTA 4422
QY 24 ----- 23
Db 4423 TLTVTPGGKKYRKLLPERKPEPEEVLKSVLKRPEEPEKPEKLEKVKKPAVPE 4482
QY 24 ----- 23
Db 4483 PPPPKPEEVEVPTVKRERKIPEPTKVPEIKPAIPLPAPEPKPPEAEVKTIKPPPEP 4542
QY 24 ----- 23
Db 4543 EPTPIAAPVTVPVVGKKAKEAKPEEAKPKGPIKGVPKKTPSPFIEARRKLRPGSGEK 4602
QY 24 ----- 23
Db 4603 PPDEAPTYQLKAVPLKFKVKEIKDIIILTESEFVGSAAIFCELYSPSTAITTWMKDGSNIR 4662
QY 24 ----- 23
Db 4663 ESPKHFRIADGDKRKLHIIDVQLSDAGEYTCVLRGLGNKEXTSTAKLVEELPVRVKTLE 4722
QY 24 ----- 23
Db 4723 EEVTVVKGOPYLISCELNERDQVVRKDKGKIVVEKPGRIVPGVIGLMRALTIINDADDTDA 4782
QY 24 ----- 23

Db 4783 GTYTTVTENANNLECSVCVKVVEVIRDMLVKPIRQHVKPKGTAFACDIARDTPNIKWF 4842
QY 24 ----- 23
Db 4843 KGYDEIPAPENDKTEILRDGNHLYLKINAMPEDIAEYAVEIEGKRYPAKLTGLGERVEL 4902
QY 24 ----- 23
Db 4903 LKPIEDVTIYEKESAFDAEISEADIPQWKLKGELLRSPPTCEIKAEGKRFLRLKVK 4962
QY 24 ----- 23
Db 4963 LDQAGEVLYQALNATITALLTVKEIELDFAVPLKDVTPERRQARFECVLTREANVWSK 5022
QY 24 ----- 23
Db 5023 GPDIIKSSDKFDIADGKKHILVINDSQDDDEGVYTAEVGKKT SARLFVIGIRLKFMS 5082
QY 24 ----- 23
Db 5083 LEDQTVKEGETATFVCELSHEKMHVWFKNDAKLHTRSRTLISSEKTHKLEMKEVTLDD 5142
QY 24 ----- FDKVAE----- 30
Db 5143 ISQIAQVKELSSSTAQLKVLEADPYTVKLDHDKTAVEKDEITLKCEVSKDVPVKWKDGE 5202
QY 31 ----- 30
Db 5203 EIVSPKYSIKADGLRLIKKADLKDKGEYVDCGTDKTKANTVVEARLIKVEKPLYG 5262
QY 31 ----- 30
Db 5263 VEVFVGETAHFEIELSEPDVHGQWKLKQPLTASPDCEIIEDGKKHILHNCOLGMRGE 5322
QY 31 ----- KLKE----- 34
Db 5323 VSFOAANAKSAANLVKELPIFTPLSDVKVFKEKDEAKFECEVSREPKTFWLKGTQEI 5382
QY 35 ----- AFMD----- 38
Db 5383 TGDDRFEIILKQTKHSMVKSAAFEDEAKYFPAEDKHTSGKLIIEGIRLKLPLKDV 5442
QY 39 ----- 38
Db 5443 AKEKESAVFTVELSHDNIRVWFKNDORLHTRTSVSMODEGKTHSITEKDLSDIDTSQIR 5502
QY 39 ----- 38
Db 5503 VEAMGMSSEAKLTVLEGDPYFTGKLQDYGVEKDEVILOCEISKADAPVKWFKDGEIKP 5562
QY 39 ----- 38
Db 5563 SKNAVIKADGKRMLILKALKSDIGQVTCDCGTDKTSGLDIEDREIKLVRPLHSVEVM 5622
QY 39 ----- 38
Db 5623 EFETARFETEISDDIHANWKLKGEALLQTPDCEIKEGKIHSVLHNCRLDQGTGGVDFQ 5682
QY 39 ----- 38
Db 5683 AANVSSAHLRVKPRVIGLLRPLKDVTVTAGEATATFCELSYEDIPVEWYLGKKLEPSD 5742
QY 39 ----- 38
Db 5743 KVVPRSEGVHTLRLRDVKLEDAGEVQLTAKDFKTHANLFVKEPPVEFTKPLEDQTVBEG 5802
QY 39 ----- KTHTC----- 43
Db 5803 ATAVLECEVSRENKVKWFKNGTEILSKSYEIVADGVRKLVTHDCTPEDIKTYTCDAR 5862
QY 44 ----- 43
Db 5863 DFKTSCNLNVPPHVEFLRPLTDLOVREKEMARFECELSRENKVKWFKDGAETKKGK 5922

QY 44 ----- 43
Db 5923 DIISKGAVRILVINKCLLDDEAEYSCVTRTARTSGMLTVLEEEAVFTKNLANIEYSETDT 5982
QY 44 ----- 43
Db 5983 IKLVCEVSPGAEVIWYKGDDEIIEITGRYEILTEGKRRLVIONAHLEDAGNYNCRPLSS 6042
QY 44 ----- 43
Db 6043 RTDGKVKVHELAAEFISKPNLEILEGEKAEFVCSISKESPPVQWRDOKTLESCKYDV 6102
QY 44 ----- 43
Db 6103 IADGKKRVLVKDATLQDMGTYYVMVGAARAAHLTVIEKLRIVVPLKDRVKEQEVVF 6162
QY 44 ----- 43
Db 6163 NCEVTEGAKAKWFRNEEAIFDSSKYIILQKDLVYTLRIRDAHLDDQANYNVSLTNRGE 6222
QY 44 ----- 43
Db 6223 NVKSAANLIVEEEDLRIVEPLKDIETMEKKSVTFWCKVNRNLNVTLKWTNGEVPFDRNV 6282
QY 44 ----- 43
Db 6283 SYRVDKVKHMLTIKDCGFPDEGEYIVTAGQDKSVAELLIIEAPTEFEVHELEDQTVTEFD 6342
QY 44 ----- 43
Db 6343 AVFSQLSREKANWYRNREIKEGKKYKPEKDGSIHRLIKDCRLDDECEYACGVEDR 6402
QY 44 ----- 43
Db 6403 KSRARLFVEEIPVEIIRPPQDILEAPGADVFLAELNKDKVEQVLRNMNMVVQDKHQM 6462
QY 44 ----- 43
Db 6463 MSEGKIHLQICDIKPRQGEYRFIAKDKEARAKLEAAPIKTADQDLVVDVGKPLTM 6522
QY 44 ----- 43
Db 6523 VVPYDAYPKABEAFKENEPISTKTIDTBTQTSFRILEAKKGDKGRYKIVLONKHGAE 6582
QY 44 ----- 43
Db 6583 GFINKLVIDVPGVPYRNLEVTETFDGEVSLAWEELTDGSGKIIGVYVVERRDIKRTWVLA 6642
QY 44 ----- 43
Db 6643 TDRAESCEFTVTGLQKGGVEYLFVRSARNRVGTGPVEDNPNVEARSKYDVPGLNVTI 6702
QY 44 ----- 43
Db 6703 TDVNRFGVSLWPEPPYDGGAEITNVYIELRDKTSIRMDTAMTVRAEDLSATVTVVBEQ 6762
QY 44 ----- 43
Db 6763 EYSFRVRAQNRIGVKGKPSAATPFVKVADPIERPSPVNLTSDDQTSQSSVLKWBPLKDG 6822
QY 44 ----- 43
Db 6823 GSPILGYIIERCEEGKDNWIRCNMKLVPELTYKVTGLEKGNKLYRVSANAGVSDPSE 6882
QY 44 ----- 43
Db 6883 ILGPLTADDAFVEPTMDLSAFKOGLEVIVPNPITILVPSTGYPRPTATWCFGDKVLETD 6942
QY 44 ----- 43
Db 6943 RVKMKTL SAYAELVISPERSDKGIYTLKLENRVKTIISGEIDVNVVIARPSAPKELKFGDI 7002

QY 44 ----- 43
Db 7003 TKDSVHLTWEPDDDDGSGPLTGYVBEKREVSRTWTVMDFVTDLEFTVPDLVQGEYLF 7062
QY 44 ----- 45
Db 7063 KVCARNKCGEPAYVDEPVNMSPTATVPDPENVKWRDRTANSIFLTWDPKNDGGSRI 7122
QY 46 ----- 47
Db 7123 KGIYVERCPRGSKWVACGEPVAKMEVTGLEEGHWYAYRVKALNRQASKPSRTEBI 7182
QY 48 ----- 52
Db 7183 QAVDTQEAPEIFLDVLLAGLTVKAGTKIELPATVTGKPEKITTKADMILKQDKRITI 7242
QY 53 ----- 58
Db 7243 ENVPKSTVTI VDSKRSDTGTIIEAVNVCGRATAVVENVLDKPGPPAAFDITDTNSES 7302
QY 59 ----- 58
Db 7303 CLLTWNPDRDDGSKITNVVERRATDSEVHKLSTVKDTNFKATKLIPNKEYIFRVA 7362
QY 59 ----- 58
Db 7363 ENMYGVGEVQASPI TAKYQFDPGPPPTRLPSDITKDAVTLTWCEPDDDDGSGPITGYW 7422
QY 59 ----- 58
Db 7423 ERLDPTDKWRCNKMVPDRTYRVKGLTNKKYFRVLAENLAGPKPSKSTEPILIKD 7482
QY 59 ----- 58
Db 7483 PIDPPMPGKPTVKDVGKTSVRLNWTKEPHDGGAKTESVIEMLKTGTDEWVRVAEGVPT 7542
QY 59 ----- 64
Db 7543 TQHLPLGMEQBYE SFRVAVNKA GSEPSDPVLCREKLYPPSPRWLEVINITKNT 7602
QY 65 ----- 73
Db 7603 ADLKWTVPEKDGSGPITNYIVEKRDVRRKGWQTVDTTVKDTCTVPTPLTEGSLYFRVAA 7662
QY 74 ----- 83
Db 7663 ENAIGSDYTEIEDSVLAKDTFTTPGPPYALAVVDVTKRHVDLKWEPPKNDGGRPIQRXV 7722
QY 84 ----- 86
Db 7723 IEKKERLGRWVKAGKTAGDCNFRVTDVIEGTEVQFQVRAENEAGVGHPSPTLSIE 7782
QY 87 DP----- 88
Db 7783 DPTSPSPPLDLHVTDAGRKHIAIAWKPEKNGGSPIIGYHVEMCPVGTCKWVRVSRPI 7842
QY 89 -EVKF----- 92
Db 7843 KDLKFVEGVDPKEYVLRVAVNAIGVSEPSISENVVAKDPCKPTIDLETHDIIVI 7902
QY 93 ----- 98
Db 7903 EGEKLSIPFRAVPVPTVSWHKDGKEVKASDRLTMKNDHISAHLEVPKSVRADAGIYTI 7962
QY 99 ----- 98
Db 7963 TLENKLGASATINVKVIGLPGCKDIKASDITKSSKLTWEPPEPDGPTILHYVLRR 8022
QY 99 ----- 98
Db 8023 EAGRRTYIPVMGENKLSWTVKDLIPNGEYFFRVKAVNVKGGEYTELKNPVTAQDPKOP 8082
QY 99 ----- 104
Db -----VEVHNA-----

Db 8083 POPPVDEVEHNPTAEAMTITWKPPLYDGGSKIMGYIIIEKIAKEERWKRCHLVPILTY 8142
QY 105 ----- 104
Db 8143 TAKGLEEGKEYOFRVRAENAAIGSEPSRATPPTKAVDPIADPKVILRTSLEVRGDEIAL 8202
QY 105 ----- 104
Db 8203 DASISGSPYTTITWIKDENVIYVPEEIKKRAAPLVRRRKGEVQEEFVPLTLQRLSIDNS 8262
QY 105 ----- 104
Db 8263 KKGESQLRVDRSLRPDGLYMIKVENDHGIAKAPCTVSLDTPGPPINFVFEDIRKTSVL 8322
QY 105 ----- 108
Db 8323 CKWEPLDDGSGSEIINYLEKKDKTKRPSSEWIVVTLRHCKYSVTKLIEGKEYLFRVA 8382
QY 109 ----- 108
Db 8383 ENRFGPGPCVSKPIVAKDPFGPPDAPDKPIVEDVTSNSMLVKWNEPKONGSPILGYWLE 8442
QY 109 ----- 108
Db 8443 KREVNSTHWSRVNKSLLNALKANVDGLLEGLTVYFRVCAENAAAGCGKFSPPDKTAHDP 8502
QY 109 ----- 108
Db 8503 ISPPGPIPRVTDTSSTTIELEWEPAPNGGEIVGYFVDKQLVGTNEWSRCTERMIKVR 8562
QY 109 ----- 108
Db 8563 QYTVKIREGADYKLRVSAVNAAGEPPGETQPTVAEPQEPAPVELDVSKGQIIMAG 8622
QY 109 ----- 108
Db 8623 KTLRIPAVVTGRPPTKVMTKEGELDKDRVIDNVGTSKSELIKDALRKDHGRYVITAT 8682
QY 109 ----- 108
Db 8683 NSCGSKFAAAARVEVDFVPGFVLDLKPVVTNRKMLLNWSDPEDDGGSEITGFIERKDAK 8742
QY 109 ----- 108
Db 8743 MHTWROPIETERSKCDITGLLEGQEKYFRVIAKNKFCGPPVEIGPILAVDPLGPPTSPE 8802
QY 109 ----- 108
Db 8803 RLTYTERTKSTITLDWKEPRNGSGSPIQGYIEKRRHDKPDPFERNKRLCPTTSFLVENL 8862
QY 109 ----- 108
Db 8863 DEHQMYEFRVAVNEIGSESESLPNVVIQDDVEVPPTIKRLSVRGDTIKVKAGEPVHIP 8922
QY 109 ----- 108
Db 8923 ADVTGLPMPKIEWSKNETVIEKPTDALQITKEEVSSEAKTELSIPKAVREDKGTYYTVA 8982
QY 109 ----- 108
Db 8983 SNRLGVSFRNVHVEYVDRPSPRNLAVTDIKAESCYLETWDAPLDNGGSEITHVIDKRDA 9042
QY 109 ----- 108
Db 9043 SRKKAEEVETNTAVEKRYGIMKLI PNQGYEFRVAVNKGYSDECKSDKVV IQDPYRLP 9102
QY 109 ----- 108
Db 9103 GPPGPKVILARTKGSMLVSWTPPLDNGGSPITGYWLEKREEGSPYWSRVSRAPITKVGLK 9162
QY 109 ----- 108
Db -----VEVHNA-----

Db 9163 GVEFNVPRLLGVKYQFRAMAINAAGIGPPSEPSDEVAGDPIFPFPPSPCEVVRKDTKS 9222
QY 109 ----- 108
Db 9223 SISLGWKPAPKDGSGPIKGIYVEMOEBGTTDMKRVNEPDKLITTCCECVVPNLKELRYRF 9282
QY 109 ----- 108
Db 9283 RVKAVNEAGESEPSDTTGEIPATDIOEPEVFIDIGAQDCLVCKAGSQRIPAVIKRPT 9342
QY 109 ----- 108
Db 9343 PKSWFEFGKAKKAMKMGVHDIPEDAQLETAENSSVILIIPECKRSHTGKYSITAKNKAQ 9402
QY 109 ----- 108
Db 9403 KTANCRVKVMYVPPPKDLKYSDIITGSCRLSWKMPDDDDGDRIGKYVIEKRTIDGKAWT 9462
QY 109 -----REEQY----- 113
Db 9463 KVNPCGSTTFVVPDLLSEQQYFFRVRAENRFGIGPPVETIQTATTARDPIYPPDPPIKLK 9522
QY 114 ----- 113
Db 9523 IGLITKNTVHLSWKPPKNDGGSPVTHYIVECLANDPTGTKEAWRQCNKRDVEELQFTVE 9582
QY 114 ----- 113
Db 9583 DLVEGGEYFRVKA VNAAGVSKSATVGPCQCRDMPSPIDLKEFEVEBEGTNVNIYAK 9642
QY 114 ----- 113
Db 9643 IKGVPFPTLTFKAPPKPNKPNKEPVLVDYTHVKNLVDDTCTLVIPOSRSDTGLYTITAV 9702
QY 114 ----- 113
Db 9703 NNIGTASKEMNLVLRGPPVGPVKFESVSADQMTLSWFPPKDDGGSKITNYVIEKREA 9762
QY 114 ----- 113
Db 9763 NRKTVHVSSEPKECTYTI PKLEGEHYFRIMAQNKYIGEPDSEPETARNLFSVPGA 9822
QY 114 ----- 113
Db 9823 PDKPTVSSVTRNSMTVNWEPEYDGGSPVTGYWLEMDTTSKRWRKRVNRDPIKAMTLGVS 9882
QY 114 ----- 113
Db 9883 KYVTGLIEGSDYQFRYAINAAGVPASLSPDATARDPIAPPPPPPKVTDWTKSSADL 9942
QY 114 ----- 113
Db 9943 EWSPLKDGSKVTGYIYVEYKEEGKEWKGDKKEVRGTLVVVTLCKEGAFYKFRVSAVN 10002
QY 114 ----- 113
Db 10003 IAGIGEPGEVTDVIEKDRLVSPDLQDASVRDRIVVHAGGVIRIIAYVSKPPPTVTWN 10062
QY 114 ----- 113
Db 10063 MNERTLPOEATIETTAISSSMVIKNCQRSHOGVYSLAKNEAGERKRTIIVDVLDPGPV 10122
QY 114 ----- 113
Db 10123 GTPFLAHLNLTWESKLTWFPEDDGGSPITNVVIEKRESDRRAWTPVTYTVTRQNAVQG 10182
QY 114 ----- 113
Db 10183 LIQKAYEFFRIAAENSIGMPFVETSEALVIREPITVPERPEDLEVKEVKTNTVLTWNP 10242
QY 114 ----- 113
Db 10243 PKYDGGSEIINVLRESRLIGTEKFKHVTNDNLLSRKYTVKGLKEGDTVEYRVSAVNIQG 10302

QY 114 ----- 113
Db 10303 GKPSFCTKPIITCKDELAPPTLHLDPRDKLTIRVGEAFALTGRYSCKPKPKVSWFKDEADV 10362
QY 114 -----NST----- 116
Db 10363 LEDDRTHIKTTPATLALEKIKAKRSDSGKYCVVVENSTGSRKGFQVNVWRPDPVGPV 10422
QY 117 ----- 116
Db 10423 SFDEVTKDYMVSWKPPDLDDGGSKITNYIIIEKKEVGKDVWMPVTSASAKTTCKVSKLLEG 10482
QY 117 ----- 116
Db 10483 KDYIFRIHAENLYGISDPLVSDSMKAKDRFRVPDAPDQPIVTEVTKDSALVTWKNPHDGG 10542
QY 117 ----- 116
Db 10543 KPITNYILEKRETMSKRWARVTKDPIHPYTKFRVPDLLLEGQCYEFRVSAENEIGIGDPS 10602
QY 117 ----- 116
Db 10603 PSKVPFAKDPIAKSPVPVNEAIDTTCNSVDLTWQPPRHDDGSKILGYIYVEYQKVGDDEW 10662
QY 117 -----YRVSV----- 122
Db 10663 RRANHTPESCPETKYKVTGLRDGQTYKFRVLAVNAAGESDPAHVPEPVLVKDRLEPPELI 10722
QY 123 ----- 122
Db 10723 LDANMAREQHKVGDTRLRLSAIRKGVPPKVTWKEDRDAPTAKRIDVTPVGSLEIRNA 10782
QY 123 -----LTV----- 125
Db 10783 AHEDGGIYSLTVENPAGSKTVSVKVLVDKPGPPRDLEVSEIRKDCSYLTWKEPLDDGGS 10842
QY 126 ----- 125
Db 10843 VITNVVVERRDVASAOWSPLSATSATSKKSHFAKHLNEGNOYLFRVAAENQYGRGPEVETPK 10902
QY 126 -----LH-----QDWL-- 131
Db 10903 PIKALDPLHPGPPKDLHHVDVKTESVLVWKNKPRDGGSPITGYLYVEYOBEGTQDWIKF 10962
QY 132 ----- 131
Db 10963 KTVNLECVVTGLOQGTKYFRVKAENIVGLPDTTPIECQEKLVPPSVVELOVKLIEG 11022
QY 132 ----- 131
Db 11023 LVVKAGTTVRFPATIRGVPVPTAKWTTDGSSEIKTDEHYTVETDNFSSVLTIKNCLRRDGT 11082
QY 132 ----- 131
Db 11083 EYQITVSNAGSKTAVHLTVLDVPGPTGPTINILDTPEHMTISWQPKDDGGSPVINY 11142
QY 132 ----- 131
Db 11143 IVEKQDTRKDTWGVVSSGSSKTLKIPHLQKGEYVFRVRAENKIGVGPPLDSTPTVAKH 11202
QY 132 ----- 131
Db 11203 KFSPPSPGKPVVTDITENAAATVSWTLPKSDGGSPITGYMYERREVTGKWVRVKNKPLAD 11262
QY 132 ----- 131
Db 11263 LKFRVTGLYEGNTYFRVFAENLAGLSKPSDPIKACRPIKPPGPINPKLKDKSRET 11322
QY 132 ----- 131
Db 11323 ADLVWTKPLSDGGSPILGYVVECOKPGTAQWNRINKDELIRQCAFVRVPLLEGNEYRPRI 11382

QY 132 ----- 131
Db 11383 KAANIVGECEPRELAESVIAKDLHPPEVELDVTCHDVTITVRVGTIRILARVKGRPEPD 11442
QY 132 ----- 131
Db 11443 ITWTBEGKVLVREKRVLDLIQDLPRVELOIKEAVRADHGKYYIISAKNSGHAQSAIVNVYL 11502
QY 132 ----- 131
Db 11503 DRPGPCONLKVNTVKENCITISWENPLDNGSGSEITNFIVEYRKPQKGHSIVASDVTKRL 11562
QY 132 ----- 131
Db 11563 IKANLIANNEYFRVCAENKVGVPITETKPTILAINPIDRPEPENLHIADKGKTFVYL 11622
QY 132 ----- 131
Db 11623 KWRRPDYDGGSPNLSYHVERRLKGDDWVERVHKSGIKETHYMWDRVCVENQIYEFVRVQTKN 11682
QY 132 ----- 131
Db 11683 EGGESDWKTEEVVVKEDLQKPVLDLKLGSVLTVRAGDTIRLEAGVYRGKPPPEVAVTKDK 11742
QY 132 ----- 131
Db 11743 DATDLTRSPRVKIDTRADSKFSLTKAKRSDGGKYVVTATNTAGSEFVAYATVNVLDKPGP 11802
QY 132 ----- 131
Db 11803 VRNLKIVDVSSDRCTVCWDPDPEDDGCETONYLEKCEKRMVMWSTYSATVLTPTGTVTR 11862
QY 132 ----- 131
Db 11863 LIEGNEYIFRVAENKIGTGPTESKPVIATKYDKPGRPDPPEVTKVSKKEETVVWNP 11922
QY 132 ----- 131
Db 11923 EYDGGKSITGYFLEKKEKHSTRWPVNVKSAIPERRMKVQNLDPDHEYQFRVAENEIGIG 11982
QY 132 ----- 131
Db 11983 EPSLPSRPVAKDPIEBPGPPTNFRVDTTKHSITLGGWKPVYDGGAPIIGYVVMRPKI 12042
QY 132 ----- 131
Db 12043 ADASPDGKRCNAAQLVKEFTVTSLDENQEYEFVCAQOVGIGRPAELKEAIKPE 12102
QY 132 ----- 131
Db 12103 ILEPPEIDLDSMRKILIVIRAGCPIRLFAIVRGRPAKVTWRKVIDNVVRKGQVDLVDT 12162
QY 132 ----- 131
Db 12163 MAFLVTPNSTRDSDGKSYTLVNPAGEKAVFVNVRLDTPGPVSDLKVSODVTKTSCHVSW 12222
QY 132 ----- 131
Db 12223 APPENDGSGQVTHYIVEKREADRKTWSTVTPVKKTSFHVTLNVPNGEYFRVTAVEYG 12282
QY 132 ----- 131
Db 12283 PGVPTDVPKPVLASDPLSEDPDRKLEVTMTKNSATLAWLPPLRDGAKIDGYITSYRE 12342
QY 132 ----- 131
Db 12343 EOQPADRWTEYSVVVDLSLVTGLKBGKKYKFRVAARNVAVGSLPREAGVYEAKEQLLP 12402
QY 132 ----- 131
Db 12403 PKILMPEQITIRAGKRLRIEAHVYGKPHPTCKWKKGEDVWTSSSHVAHVHADSSILIIK 12462
QY 132 ----- 131

Db 12463 DVTRKDSGYISLTAENSSGTDQKIKVVVMDAPGPPQPPFDISDIDADACSLSWHIPLD 12522
QY 132 ----- 131
Db 12523 GGSNITNYIVEKCDVSRGDWVTALASVTYTSRVRGKLIPGOEYIFRVAENRPGISEPLT 12582
QY 132 ----- 131
Db 12583 SPKMAQFPFGVPSEPKNARVTKVNKDCIFVAWDRPDSDDGSPILIGYLIERKERNLLWV 12642
QY 132 ----- 131
Db 12643 KANDTLVRSTEYPCAGLVEGLEYSFRYIALNKAGSSPPSKPTTEYVTARMPVDPGKPEVI 12702
QY 132 ----- 131
Db 12703 DVTKSTVSLIWARPKHDGSKIIIGYFVEACKLPGDKWVRCNTAPHQIPOEYATATGLEEK 12762
QY 132 ----- 131
Db 12763 AOYQFRAIARTAVNISPPSESPDVTILAEENVPPRIDLSVAMKSLTLVKAGTNVCLDATV 12822
QY 132 ----- 131
Db 12823 FGKPMPTVSWKKDGTLLKPAEGIKMAMQRLCTLEFSVNRKDSGDYTTITAENSSGSKSA 12882
QY 132 ----- 131
Db 12883 TIKLVLDKPGPPASVKINKMYSDRAMLSWEPPLEDGSGSEITNYIVDKRETSRPNQAQVS 12942
QY 132 ----- 131
Db 12943 ATWPTSCSVEKLEHGEYQFRICAENKYGVDVTEPAIAKNPYDPGRCDDPPVISNI 13002
QY 132 ----- 131
Db 13003 TKDHMTVSNKPPADGGSPITGYLLEKRETOAVNWTKVNRKPIIERTLAKTLOEGTEYE 13062
QY 132 ----- 131
Db 13063 FRVTAINKAGCPKPSDASKAAAYARDPOYPAPPAPFKVYDTRSSVSLSWGKPAYDGGSP 13122
QY 132 ----- 131
Db 13123 IIGLYLVKRADSDNWVRNLPONLQKTRFEVTGLMEDTQYQFRVAVNKGIGYSDPSDVP 13182
QY 132 ----- 131
Db 13183 DKHYPKDILIPPEGELDADLRKTLILRAGVTMRLYVPVKGRRPPPKITWSKPNVNLDRIG 13242
QY 132 ----- 131
Db 13243 LDIKSTDFDTFLRCENNVKYDAGKYILTLNENSCGKKEYTIVVKVLDTPGPPVNVTVKEIS 13302
QY 132 ----- 131
Db 13303 KDSAYVTWEPPIIDGSGPIINYVVKRDAERKSWSTVTTECSKTSPFRVANLEEGKSYFFR 13362
QY 132 ----- 131
Db 13363 VFAENEYIGDPGETRDAVKASQTPGVVDLKVRSVSKSSCSGTGKKPHSDGSGRIIGYV 13422
QY 132 ----- 136
Db 13423 VDFLTEENKQWRVMSLSLOYSAKDLTEGKEYTFRVSAENEGEPTSEITVVVARDVVA 13482
QY 137 ----- 136
Db 13483 PDLDLKGLPDLCLAKENSFRLLKIPKGPAPSVWKKGEDPLATDTRVSESSAVNTT 13542
QY 137 ----- 136

Db 13543 LIVYDCQSDAGKYTITLKNVAGTKBGTISIKVVGKPGIPTGPIKFDDEVTAEMTLKWP 13602
Qy 137 ----- 136
Db 13603 PKDDGSEITNYILLEKRDVNKNWTCASAVQKTTFRVRLHEGMEYTFRVAENKYGVG 13662
Qy 137 ----- 136
Db 13663 EGLKSEPIVARHPFDVPOAPPPNIVDVHRDVSLSLTWTDPKKTKGSGPITGYHLEFKERNS 13722
Qy 137 ----- 136
Db 13723 LLWKRANKTPIMRDKFKVTGLTEGLEVEFRVMAINLAVGKPSLPSEPVVALDPIDPPGK 13782
Qy 137 ----- 136
Db 13783 PEVINITRNSVTLIWTPEPKYDGGHKLTCYIIVEKRDLPKSKWMKANHVNPCEAFVTVDLV 13842
Qy 137 ----- 136
Db 13843 EGGKYEFRIRAKNTAGAISAPSESTETIICKDEYEAPTIVLDPTIKDGLTIKAGDTIVLN 13902
Qy 137 ----- 136
Db 13903 AISILGKPLPKSWSKAGKDIRPSDITQITSTPTSSMLTIKYATRKDAGEYTIATNPFG 13962
Qy 137 ----- 136
Db 13963 TKVEHVKVTVLDVPCPPVPEISNVSAEKATLTWTPLEDGSGPIKSYILEKRETSRLW 14022
Qy 137 ----- 136
Db 14023 TWSIEDIQSCRHVATKLIQNEGYFRVSAVNHYGKGPVQSEPVKMVDRFPGPPGPEKPE 14082
Qy 137 -----KCVSN----- 142
Db 14083 VSNVTKNTATVSWKRVPDDGSEITGYHVERREKSLRWRAIKTPVSDLRCKVTGLQEG 14142
Qy 143 -----:---: 142
Db 14143 STYEFVSAENRAGIGPPSEASDVLMDAAYPGPPSPNHVDTTKKSASLANGKPHYD 14202
Qy 143 ----- 142
Db 14203 GGLEITGVVHEQKVGDEAWIKDTGTALRITQFVVPDLOTKEYNFRISAINDAGVGP 14262
Qy 143 -----KALPAP----- 148
Db 14263 AVIPDVEIVEREMAPDFELDAELRRTLVRAGLSIRIFVPKGRPAPEVVTWKDNINLKN 14322
Qy 149 ----- 148
Db 14323 RANIENTESFLLIIPECNRYDTCGFVMTENPAGKSGFVNVVRVLDTPGPVNLNRPDI 14382
Qy 149 -----IEK----- 151
Db 14383 TKDSVTLHWDLPDLIDGGSRTITNYIVEKREATKSYSTATTKCHKCTYKVTGLSGCEYFF 14442
Qy 152 ----- 151
Db 14443 RVMAENEYGIGEPTTETPEPVKASEAPSPDPSLNIMDITKSTVSLAWPKPKHDGSKITGY 14502
Qy 152 ----- 151
Db 14503 VIEAQKGSQDWHITTVKGLECVVRNLTEGEYTFQVMAVNSAGRSAPRESRPVIVKEQ 14562
Qy 152 ----- 151
Db 14563 TMLPELDLGIYQKLVIAKAGDNIKEIPLGRPKPTVTKWKGDIKQIQRVNFETTAT 14622
Qy 152 ----- 151
Db 14623 STILNINECVRSDSGPYPLFARNIVGEVDVITIQVHDIQPGPTGPIKFDEVSSDFVTF 14682

Qy 152 ----- 151
Db 14683 WDPENDGGVPISNVVVEMROTDSTTWELATTIVRTTYKATRLTTGLEQYFRVKAQNR 14742
Qy 152 ----- 151
Db 14743 GVGPGITSACIVANYPFKVPGPPTQVAVTKDSMTISWHEPLSDGSGPILGYHVERKE 14802
Qy 152 -----TISKA----- 156
Db 14803 RNLGMOTVSKALVDPGNIFKSSGLTDGIAYEFRVIAENMACKSPSPSEPMALADPIDP 14862
Qy 157 -----:---: 156
Db 14863 PGKVPPLNITRHTVTLAKWAKPEYTGKPKITSYIIVEKRDLPNGRWLKNFNSILENEFTVS 14922
Qy 157 ----- 156
Db 14923 GLTEDAAVEFRVIAKNAAGAISPPSPSDAITCRDDVEAPKIKVDVKFKDVTILKAGEAF 14982
Qy 157 ----- 156
Db 14983 RLEADVSGRPPPTMEWSKDKKELEGTAKEIKIADFSTNLVKNKSTRRDSGAYTLTATNP 15042
Qy 157 ----- 156
Db 15043 GGFAKHIFNVKVLDRPGPPGPEGLAVTEVSEKCVLSWFPPLDDGGAKIDHIVIVOKRETSR 15102
Qy 157 ----- 156
Db 15103 LAMTVASEVQVTKLVTKLLKQNEYIFRVMAVKNYGVGEPLSEPVLA VNPYGPDPKK 15162
Qy 157 ----- 156
Db 15163 NPEVTTITKDSMVVWCWHPDSDGSEIINYIVERRRKAGORWIKCNKKTTLTDLRYKVSGL 15222
Qy 157 ----- 156
Db 15223 TEGHEYEFIRMAENAGISAPSPSPFYKACDTPFKPGPGNPRVLDTSRSSISIANKKP 15282
Qy 157 ----- 156
Db 15283 IYDGSSEITGYMVEIALPEDEWQIVTPPAGLKATSYITIGLTENQEYKIRIYAMNSEGL 15342
Qy 157 ----- 156
Db 15343 GEPALVPGTPKAEDRMPLPPEIELDADLRKVVTIRACCTLRLFPVPIKGRPAPEVKWARDHG 15402
Qy 157 ----- 156
Db 15403 ESLDKASIESTSSYTLIIVGNVNRFDGSKYILITVENSNGSKSAFVNVVRVLDTPGPPQDLK 15462
Qy 157 ----- 156
Db 15463 VKEVTKTSVTLTWDPPLLDGSGKIKNYIVEKRESTRKAYSTVATNCHKTSWKVDLOEGC 15522
Qy 157 ----- 156
Db 15523 SYFVRVLAENEYIGLPAETAESVKASERPLPGKITLMDVTRNSVLSWEKPEHDGSR 15582
Qy 157 ----- 156
Db 15583 ILGYIVEMQTKSGDKWATCATVKTVEATITGLIGEEYSFRVSAQNEKIGSDPQLSPVP 15642
Qy 157 ----- 156
Db 15643 IAKDLVTPPAFKLILFNFTVVLAGEDLKVDVFPFICRPTPAVTHKDNVPLKQTTTRVNAEST 15702
Qy 157 ----- 156
Db 15703 ENNSLLTIKDACREDVGHYVVVVKLTNSAGEALETNLVLDKPGPTGPKVMDVETADSIT 15762

QY	157	-----	156
Db	15763	LSWGPPIYDGGSSINNYIVEKRDTSPTTQIIVSATVARTTIKACRLTKCEYQFRIAAEN	15822
QY	157	-----	156
Db	15823	RYGKSTYLNSEPTVAQYPFKVPFGPTPVYTLSSRDSMEVQWNEPISDGGSRVIGYHLER	15882
QY	157	-----	156
Db	15883	KERNLSILWKLANKTPIQTKFKTKTGLEGEYEFVRVSAENIVGIGPKSVKSECYVARDCP	15942
QY	157	-----	156
Db	15943	DPPGRPEAIIVTNSVTLQWKKPTYDGGSKIITGYIVEKKELPEGRWMAKASFTNIIDTHFE	16002
QY	157	-----	156
Db	16003	VTGLVEDHRYEFVRVIAARNAAGVSEFSESTGAITARDEYDPPRISMDPKYKDTIVVHAGE	16062
QY	157	-----	156
Db	16063	SPKYVDADIYCKPIPTIOWIKGQDLSNTARLEIKSTDFATLSVKDAVRVDSGNYILKAK	16122
QY	157	-----	156
Db	16123	NVAGERSVTNVKVLDRPPEGPVVISGVTAEKCTLAWKPPLODGGSDIINIVIVERRET	16182
QY	157	-----	156
Db	16183	SRLVMTVDANVQTLSCVKYTKLLEGNEYTRIMAVNKGVGEPLESEPVVAKNPVVPA	16242
QY	157	-----	156
Db	16243	PKAPEVTITKDSMIVVWRPASDGGSEILGYVLEKRDKEGINWTRCHRLIGELRLVT	16302
QY	157	-----	156
Db	16303	GLIENHDYEFVRVSAENAGLSEPPSAQKACDPIYKPGPNPKVIDITRSSVFLWS	16362
QY	157	-----	156
Db	16363	KPIYDGGCEIOGYIVEKCDVSGWMTMCTPTGINKNTINEVEKLEKHEYNFRICAINKA	16422
QY	157	-----	156
Db	16423	GVGEHADVPPIIVEKLEAPDIDLDELRLKIINIRAGGSLRLFVPIKGRPTPEVKGKV	16482
QY	157	-----	156
Db	16483	DGEIRDAAIIDVTSSFTSLVDNVNRYDSKVTLTLENSSGTKSAFVTVRVLDTPSPVYN	16542
QY	157	-----	156
Db	16543	LKVTETKDSVITWBPPLLDGGSKIKNYIVEKREATRKSAAVVTNCHKNSWKIDQLOE	16602
QY	157	-----	156
Db	16603	GCSYFVRVTAENEYIGLPAQTADPIKVAEVPQPGKITVDVTRNSVLSWTKPEHDGG	16662
QY	157	-----	156
Db	16663	SKIIQIVEMOAKHSEKSECARVKSLOAVITNLTQGEYLFVRVAVNEKGRSDPRSLAV	16722
QY	157	-----	156
Db	16723	PIVAKDLVIEPDVKPAFSSYSVQGDLEKIEVPISGRPKPTITWTKDGLPLKQTRINVT	16782
QY	157	-----	156
Db	16783	DSLDLTTLISIKETHKDDGGQGITVANVVGQKTASIEIIVTLDKDPKPGVVKFDDVSAES	16842
QY	157	-----	156

Db	16843	ITLSWNPPLYTSGCOITNYIVOKRDTTTTWDVVVSATVARTTLKVKLTKGTGEYQFRIFA	16902
QY	157	-----	156
Db	16903	ENRYGOSFALESPIVAQYPIYKEPPTFPFATAISKDSMVIQWHEPNNGGSPVIGYHL	16962
QY	157	-----	156
Db	16963	ERKERNLSILWTKVNTIIHDTQFKAQNLBEGIEYEFVRVVAENIVGVKASKNSECYVARD	17022
QY	157	-----	156
Db	17023	PCDPPCTPEPIMVKRNEITLQWTKPVYDGGSMITGVIVEKRDLPDGRWMAKASFTNVIETQ	17082
QY	157	-----	156
Db	17083	FTVSGLTEDQRYEFVRVIAKNAAGAIKSPSDSTGPITAKDEVELPRISMDFKFRDTIVVNA	17142
QY	157	-----	156
Db	17143	GETFRLEADYHGKPLPTIEWLRGDKIEBSARCEIKNTDFKALLIVKDAIRIDGGQYILR	17202
QY	157	-----	156
Db	17203	ASNAGSKSPVNVKVLDRPPEGPVQVTGVTSEKSLTWSPLQDGGSDISHYVVEKR	17262
QY	157	-----	156
Db	17263	ETSRLAWTVVASEVVTNSLKVTKLLEGNEYVFRIMAVNKGVGEPLESAFVLMKNPFVLP	17322
QY	157	-----	156
Db	17323	GPPKSLVETNIAKDSMTVCNRPDSDGGSEIIGYIVEKRDORSIRWIKCNKRITDLRLR	17382
QY	157	-----	156
Db	17383	VTGLTEDHEYEFVRVSAENAGVGEPSPATVYKACDPVKFPGPTNAHIVDTTKNSITLA	17442
QY	157	-----	156
Db	17443	WKPPIYDGGSEILGYVVEICKADEEEMQIVTQOTGLRVTRFEISKUTEHQEYKIRVCALN	17502
QY	157	-----	156
Db	17503	KVGLGEATSVPGITVKPEDKLEAPELDLDSFLRKGIVVRAGGSARIHIPFKGRPTPEITWS	17562
QY	157	-----	156
Db	17563	REEGEFTDKVQIEKGVNYTQLSIDNCDNRNDAGKYILKLENSGSKSAFVTVKVLDTPGPP	17622
QY	157	-----	156
Db	17623	QNLAVKEVRKDSAFVWPEPPIIDGGAKVKNYVIDKRESTRKAYANVSSCKSKTSFKVENL	17682
QY	157	-----	156
Db	17683	TEGAIYYFRVMAENEFVGVPVETVDVAKAAEPPSPGKVTLTDSQTSASLAWKEPHEHD	17742
QY	157	-----	156
Db	17743	GGSRVLGYVVEMQPKTEKWSIVAESKVCNAVVTGLSSGOEYQFRVKAYNEKGS DPRVL	17802
QY	157	-----	156
Db	17803	GVPTIAKDLTIQPSLKLPTNTYSIQAGEDLKIEIPVIGRPRPNISWVKOGEPLKQTRVN	17862
QY	157	-----	156
Db	17863	VEETATSTVLHIKEGKNDDFGKYTVTATNSAGTATENLSVIVLEKPGPPVGRFDEVSA	17922
QY	157	-----	156

Db 17923 DFVWISWEPAYTGCQISNYIVEKRDTTTTTHWMSATVARTTIKITKLKTGTQYQPRI 17982
QY 157 ----- 156
Db 17983 FAENRYGKSAPLDSKAVIVQYFKEPGPGPFVTSISKDQMLVQWHEPVNDGGTKIIGY 18042
QY 157 ----- 156
Db 18043 HLEQEKNSILWVLKLNKTPIDQTKFKTGLDGLGLEYEFKSAENIVGIGKPSKSECFA 18102
QY 157 ----- 156
Db 18103 RDCPPGRPEAIVITRNNTLKWKKPAYDGGKITGYIVEKKDLPDGRWMKASFTNVLE 18162
QY 157 ----- 156
Db 18163 TEFTVSLVEDORYEFVRIARNAAGNFSEPSDSSGAIYARDEIDAPNASLOPKYKDVIV 18222
QY 157 ----- 156
Db 18223 HAGETVLEADIRGKPIPDVVWSKDGKELETAARMEIKSTIQKTTLVVKDCIRTDGGQY 18282
QY 157 ----- 156
Db 18283 ILKLSNVGCKSIPITVKVLDORPGPEGLKVTGVTAEKCYLAWNPPLQDGGANISHYII 18342
QY 157 ----- 156
Db 18343 EKRETSRLSWQVSTEVOALNYKVKLLPGNEYIFRVMVKNKYGIGEPLESGPVACNPY 18402
QY 157 ----- 156
Db 18403 KPPGPPSTPEVSAITKDSMVVTWARPVDDGTEIEGYILEKRDKEGVWTKCNKKTLDL 18462
QY 157 -----KGQPREPOVY----- 166
|:| | | |
Db 18463 RLVRTGLTEGHSYEFVRAAENAGVGESEPSVFYACDALYPPGPPSNPKVTDTSRSV 18522
QY 167 ----- 166
Db 18523 SLAWSKPIYDGGAPVKGYVVEKEAAADENWTCPTTGLQKQFTVTKLKENYENFRIC 18582
QY 167 -----TLP----- 169
| | |
Db 18583 AINSEGVGEPATLPGSVVAQERIEPPEIELDALRKVVVLASATLRLFTVTKGRPEPEV 18642
QY 170 ----- 169
Db 18643 KWKAEGILTDRAQIEVTSFTMLVIDNVTRFDSGRYNLTLENNSGSKTAEFVNVRLDSP 18702
QY 170 ----- 169
Db 18703 SAPVNLITIREVKDSVTLISWEPLLDGGAKITNYIVEKRETRKAYATITNNCTKTPRI 18762
QY 170 ----- 169
Db 18763 ENLQEGCSYFRVLASNEYIGLPAETPEPVKVSPEPLPPGVRVLVDVTRNTATIKWKP 18822
QY 170 ----- 169
Db 18823 ESDGSKITGYVEMQYKGEKSWCTQVKTLEATISGLTAGEEYVFRVAANVEKGRSDP 18882
QY 170 ----- 169
Db 18883 RQLGVPVIARDIEIKPSVELPFHTFNVKAREQLKIDVPFKRQONTVNWVKDQTLKETT 18942
QY 170 ----- 169
Db 18943 RVNVSSKVTVTSLSIKEASKEDVGTYELCVSNSAGSITVPTITIVLDRPGPGPIRIDEV 19002
QY 170 ----- 169
Db 19003 SCDSITISWNPPEYDGGCQISNYIVEKETTSTTWHIVSQAVARTSIKIVRLTTGSEYQF 19062

QY 170 ----- 169
Db 19063 RVCAENRYCKSSYSESSAVAEYFPPPGPCTPKVHATKSTMLVTWQVPVNDGSRVI 19122
QY 170 ----- 169
Db 19123 GYHLEYKERSSILWSKANKILIADTQMKVSGLDEGLMEYRYVYAENIAGICKSKSCBPV 19182
QY 170 PSRD----- 173
|:| | |
Db 19183 PARDFCDPPGQPEVTNITRKSVLSKWSKPHYDGGAKITGYIVERRELDPGRWLKCNVTNI 19242
QY 174 ----- 173
Db 19243 QETYFEVTELTEDQRYEFVRFARNADSVSEPSSESTGPIIVKDDVEPPRVMDVKFRDVI 19302
QY 174 ----- 173
Db 19303 VVKAGEVLKINADIAGRPLPVISWAKDGEIEERARTEIISTDNHTLLTVKDCIRRDTCQ 19362
QY 174 ----- 173
Db 19363 YVLTILKNVAGTRSAVNCKVLDKFPAGPLLEINGLTAEKCSLWGRPQEDGADIDYI 19422
QY 174 ----- 173
Db 19423 VEKRETSHLAWTICEGELQMTSCKVTKLLKGYEYIFRVTGVNKGVGGEPLSVAILKALDP 19482
QY 174 ----- 173
Db 19483 FTVPSPTSLTITSVTRESMTLWCSRPESDGSEISGYIIERREKNSLRWVRNKKPVYD 19542
QY 174 ----- 173
Db 19543 LRKSTGLREGCEYERYVYAENAGLSLSETSPILRAEDPVFLPSPPSKPKIVDSGKTT 19602
QY 174 ----- 173
Db 19603 ITIAWVPLFDGGAPITGYTVEYKKSDDTDMKTSIQSLRGTEYITISGLTTGAEYVFRVKS 19662
QY 174 ----- 173
Db 19663 VNVGASDPDSQPIAKEREPEPLFDIDSEMRKTLIVKAGASFTMTVPERGRPVNVL 19722
QY 174 ----- 173
Db 19723 WSKPDTDLRTRAYVDTTDSRTSLTIENANRNDSGKYTLTIQNVLSAASLTLLVVKVLDTPG 19782
QY 174 ----- 173
Db 19783 PPTNITVQDVTKESAVLSWDVPENDGGAPVKNYHIEKREASKAWSVYNNCNRLSYKVT 19842
QY 174 ----- 173
Db 19843 NLOEAIYFRVSGENEFVGIAPAEKTEGVKITEKPSPEKLGVTISIKDSVSLTWLKE 19902
QY 174 ----- 173
Db 19903 HDGSRIVHYVVEALEKGQKNWVKCAVAKSTHHVVSGLRENSEYFFRVFAENQAGLSOPR 19962
QY 174 ----- 173
Db 19963 ELLLPVLKEOLEPPEIDMKNFPSHTVYVRAGSNLKVDPISGKPLPKVTLSDRGVPLKA 20022
QY 174 ----- 173
Db 20023 TMRNTEITAENLTINLKESVTADAGRYEITAANSSGTTKAFINIVVLDROPPTGPVVI 20082
QY 174 ----- 173
Db 20083 SDITEESVTLKWEPPKYDGGSQVTNYILLKRETSVAVTEVSATVARTMMKMLTTGEE 20142

QY 174 ----- 173
Db 20143 YQFRIKAENRFGISDHIDSACVTKLPYTPPGPSTPWNTVNTRESITVGHPEVSNNGS 20202
QY 174 ----- 173
Db 20203 AVVGYHLEMKDRNSILWQKANKLVIRTHPKVTTISAGLIIEFRVYAENAAAGVKPSHPS 20262
QY 174 ----- 182
 :---ELTKNOVSL-
 :::| | |
Db 20263 EPVLAIDACEPPNRVTRITDISKNSVLSWQPAFDGSGKITGYIVERDLPDGRWTKASF 20322
QY 183 ----- 185
 -TCL-
 | | |
Db 20323 TNVTETQFIISGLTONSOYEFVRFARNAGVSGISNPSEVGPITCIDSYGGPVIDLPLEYT 20382
QY 186 ----- 185
Db 20383 EVVKYRAGTSVKLRAGISGKPAFTIEWYKDDKELOTNALVCVENTTDLASILIKDADRLN 20442
QY 186 ----- 185
Db 20443 SGCYELKLRNAGSASATIRVQILDKRPGPGPIEFKVTVAEKITLLWRPPADDGGAKIT 20502
QY 186 ----- 185
Db 20503 HYIVEKRETSRVVWSMVSEHLEBCIITTKIIKNEVIFRVRANKYIGIGEPLESDSVA 20562
QY 186 ----- 185
Db 20563 KNAFVTPGPGPIPEVTKITKNSMTVYVWSRPIADGGSDISGYFLEKRDKKSLSGWFVKLKET 20622
QY 186 ----- 185
Db 20623 IRDTRQKVTGLTENSQYRVCANRAGQGFSEPSEFYKAADPIDPPGPPAKIRIADST 20682
QY 186 ----- 185
Db 20683 KSSITLGSKPVDYDGSAAVTGYVVEIRQEGBEWTTVSTKGEVTRTEYVYVSNLKPGVNY 20742
QY 186 ----- 185
Db 20743 FRVSANACOGGPIEMNEPVQAKIDLEAPEIDLVALRTSVIAKAGEDVQVLIPEKGRP 20802
QY 186 ----- 185
Db 20803 PPVTVWRKDEKNLGSRDARYSIENTDSSLLTIPQVTRNDTGKYLITTIENGVGPEKSTVS 20862
QY 186 ----- 185
Db 20863 VKVLDTPAACQKLQVHVSRGVTLLWDPLIDGGSPINNVIEKRDATKRTWSVYSHKC 20922
QY 186 ----- 185
Db 20923 SSTSKLIDLSEKTPFRVLAENEIGEPCEPTEPVKAAEVPAPIRDLMSKDSTKTSV 20982
QY 186 ----- 185
Db 20983 ILSWTRPDFDGGSVITEYVYVVERKKGQETWSHAGISKTCEIEVSQLEQSVLEFRVFAKN 21042
QY 186 ----- 185
Db 21043 EKGLSDPVTGPIVTKELIITPEVDLSDIPGAQVTVRIGHNVHLELPYKPKPSISWLK 21102
QY 186 ----- 185
Db 21103 DGLPKSEFVRSKTEKNKITLSIKNAKHBGGKYTVILDNAVCRVIAPIVITVTLGPPSK 21162
QY 186 ----- 185
Db 21163 PKGPIRDEIKADSVILSDVPEDNGGGETCYISIEKRETSQTNWKMVCSSVARTPKVP 21222
QY 186 ----- 185

Db 21223 NLVKADEYQFRVRAENRYGVSQLPVSSIIIVAKHQFRIPGPKPVYINVTSDGMSLTWDA 21282
QY 186 ----- 190
 -VKGFY-
 | | |
Db 21283 PVDGGSSEVTGPHVEKKEKERNISILWQKNTSPISGREYRATGLVEGLDYQFRVYAENSAGL 21342
QY 191 ----- 193
 --PSD-
 | | |
Db 21343 SSFSDPSKFTLAVSPVDPGTPDYIDVTRITITLKNPPLDDGSGKIVGYSIEKROGNER 21402
QY 194 ----- 193
Db 21403 WVRCNFTDSECOYTVTGLSPGDRYEFRIIARNAVGTISPPSQSGSIIMTRDENVPPIVE 21462
QY 194 ----- 193
Db 21463 FGPEYFGLIIKSGESLRIKALVOGRVPVPTWFKDGVETIEKRMMEITDVLGSTSIFVR 21522
QY 194 ----- 193
Db 21523 DATRDHRGVTVVEAKNAGSNAKAEIKVKVQDTPGKVVGPITFTNITGKMTLWWDAPLND 21582
QY 194 ----- 193
Db 21583 GCAPITHYIIIEKRETSRLAWALIEDKCEAQSYTAIKLINGNEYQFRVSANVKFGVGRPLD 21642
QY 194 ----- 193
Db 21643 SDPVVAIOYTVDPADGPIPEPSNITGNSITLTWARPESDGGSEIQOYILERREKKSRTWV 21702
QY 194 ----- 193
Db 21703 KVISKRPISRETRKVTGLTEGNEYEFHVMAENAGVGPASGISRLIKREPVNPBGPTV 21762
QY 194 ----- 198
 ---IAYEW-
 :::| | |
Db 21763 VKVTDTSKTTVSLWSKPVDFDGGMEIIGIIECMCKADLGDWHKVNAEACVKTRTVTDLQ 21822
QY 199 ----- 198
Db 21823 AGEYKFRVSAINGAGKGDSCVETGTIKAVDRLTAPELIDANFKQTHVVRAGASIRLFI 21882
QY 199 ----- 201
 ---ESN-
 | | |
Db 21883 AYQGRPTTAVWSKPSDNLRLADIHTDSEFTLTVCNCRNDACKYTLTVENNSGSKSI 21942
QY 202 ----- 201
Db 21943 TFTVKVLDTPGPPGPIITFKDVTGRSATLMDAPLLDGGARITHYVVEKREASRSQWVIS 22002
QY 202 ----- 201
Db 22003 EKCTROIKFVNDLAEGVYYFRVSANVEYGVGEPEMPEPIVATEQPAPRRLDVDTSK 22062
QY 202 ----- 201
Db 22063 SSAVLAWLKPDHDDGSRITGYLLEMRKSGDFWVEAGHTKOLTFTVERLVEKTEYEFVK 22122
QY 202 ----- 201
Db 22123 AKNDAGYSEPREAFSSVIIKEPQIEPTADLTGINTQLITCKAGSPFTIDVPISGRPAKV 22182
QY 202 ----- 201
Db 22183 TWKLEMLRKETDRVSITTTTKDRTTLAVKDSMRGDSGRYFLTLENTAGVKTFSVTVVIG 22242
QY 202 ----- 204
 ---CQP-
 | | |
Db 22243 RGPVTPGPIEVSSVSAESCVLSWGEKPKDGGGTETITNIVKRESGTTAWQLVNSVKRTQ 22302
QY 205 ----- 204

Db 22303 IKVTHLTMYEYFRVSSNREFGSKPLESAPIAIEHPFVPPSAPTRPEVYHVSNAMSI 22362
QY 205 -----ENNYKTT----- 211
Db 22363 RWEPPYHGGSKIIGVWEKKERNITLWVKNKVCLECNKYKVTGLVGELEYQFRTYALN 22422
QY 212 ----- 211
Db 22423 AAGVSKASEASRPIMAQNPDAPGRPEVTDVTRSTVSLWSAPAYDGGSKVVGIIERKP 22482
QY 212 ----- 211
Db 22483 VSEVGDGRWLKCNITYTSDNFFVTALSEGDTYEFRLAKNAAGVISKSESTGPVTCRD 22542
QY 212 --PP----- 213
Db 22543 EYAPPAELDARLHGLDVTIRAGSDLVDAVGGKPEPKIITWKGDKELDICEKVSLOVT 22602
QY 214 -----VLDSG----- 219
Db 22603 GKRAVIAKFCDRSDSGKYTLTKNASGTKAVSVHVKVLDSPGCKLTVSRVTOEKCTL 22662
QY 220 ----- 219
Db 22663 AWSLPQEDGGAETHYIIVERRETSRLNMVIVEGECPTLSYVVTRLIKNNEYIFRVRVANK 22722
QY 220 -----SFFL----- 223
Db 22723 YGPGVPESEPIVARNSTIIPSPGPPEVGTGKEHIIQWTKPESDGNESINLVDKR 22782
QY 224 ----- 223
Db 22783 EKKSLRWTRVNDKYVYDTRLKVTSLMEGCDYQFRVTA^NAGNSEPSEASNFISCREPS 22842
QY 224 ----- 223
Db 22843 YTPGPPSAPRVVDTKHSISLAWTKPMYDGGTDIVGYVLEMQEKDQDQWYRHTNATIRN 22902
QY 224 ----- 223
Db 22903 TEFTVPLKMGQKYSFRAAVNVKMGSEYSESIAEIEPVERIEIPDLELADDLKKTVTIR 22962
QY 224 -----YSKLTVDK----- 231
Db 22963 AGASLRLMVSVSRPPPVITWSKOGIDLASRAIIDTTESYSLLIVDKVNRVDAGRYTIEA 23022
QY 232 ----- 231
Db 23023 ENQSGKKSATVLVKYVDTPGCPSPVKVEVSRDSVTITWEIPTIDGAPVNNYIVEKREA 23082
QY 232 ----- 231
Db 23083 AMRAFKTVTCKSKTLRISGLVEGTMYYFVRLPENIYIGEPCESTDVAVLSEVPLPA 23142
QY 232 ----- 231
Db 23143 KLEVVDVTKSTVLAWEKPLYDGGSLRAGYLEACKAGTERMKVYVTLKPTVLEHTVTSL 23202
QY 232 ----- 231
Db 23203 NEGEQYLFRIRAEQNEKGVSEPRETVTAVTVQDLRLVLPDIDLSMPQKTIHVPAGRPVELV 23262
QY 232 ----- 231
Db 23263 IPIAGRPPPAASWFFAGSKLRESERVVETHTKVAKLTIRETTIRDTGEYTLKKNVTGT 23322
QY 232 ----- 231
Db 23323 TSETIKVIIIDKPGPTGPIKIDEIDATSIITISWPELDGGAPLSGYVVEQDRAHPGW 23382
QY 232 ----- 231
Db 23383 LPVSESVTRSTFKFTRLLEGNEYVFRVAATNRFNGIGSYLQSEVIECRSSIRIPGPPETLQ 23442

QY 232 ----- 231
Db 23443 IPDVSRDGHTLTWYPPEDDDGGSQVTGYIVERKEVRADRRVRVKNKVPVMTTRIRSTGLTEG 23502
QY 232 ----- 231
Db 23503 LEYEHRTVAINARGSGKPSRPSKPIVAMDPIAPPCKPQNPRVTDTRTTSVSLAWSVPDE 23562
QY 232 ----- 231
Db 23563 GGSKVTVGLIEMQKVDQHEWTKCNTPTTKIREYTLTHLPQGAERYFRVLACNAGGPPEA 23622
QY 232 -----SRWOG----- 237
Db 23623 EVPGTVKVTEMLEYPDYELDERYQEGIFVRQGGVIRLTIPIKKGKPPICKWKTEGQDISK 23682
QY 238 ----- 237
Db 23683 RAMIATSEHTELVIKEADRGDSGTDLVLENKCGKAVYIKVRVIGSPNSPEGPLEYDD 23742
QY 238 ----- 237
Db 23743 IOVRSVRVSWRPPADDDGADILGYILERREVPKAAWYITDSRVGTSLTVVKGLENVEYH 23802
QY 238 ----- 237
Db 23803 FRVSAENQFGISKPLKSEPVTPKTLNPPPEPPSNPPEVLDVTKSSVLSWSRPKDDGGS 23862
QY 238 ----- 237
Db 23863 RVTGYIIERKETSTDKWVRHNKQITTTMYTGLVLPDAEYQFRIIAONDVGLSETSPAS 23922
QY 238 ----- 237
Db 23923 EPVCKDPDKPSQPCLEILISKDSVTLQWEKPECDDGKILGYWYVEYRQSGDSANKK 23982
QY 238 ----- 237
Db 23983 SNKERIKDKQFTIGCLEATEYFRVFAENETGLSRPRRTAMSIKLTSGEAPGIRKEM 24042
QY 238 ----- 237
Db 24043 KDVTTKLGAAQLSCQIVGRPLPIKWYRFGKELIQSRKYMSSDGRHTLTVMTEQED 24102
QY 238 ----- 237
Db 24103 EGVYTCIATNEVGEVETSSKLLLOATPOFHPGYPKLEKYGAVGSTLRLHVMYIGRPVPA 24162
QY 238 -----NVF----- 240
Db 24163 MTFHGOQLLQNSENITIENTEYTHLVKMNVRKTHAGKYKQLSNVFGTVDAILDVEI 24222
QY 241 ----- 240
Db 24223 QDKPKDPTGPIVIEALLKNSAVISWKPADDGGSWITNYVVEKCEKAEGAEHQLVSSAIS 24282
QY 241 ----- 240
Db 24283 VTTCRIVNLTENAGYYFRVSAQNTFGISDPLEVSSVVIKSPFEKPGAPKPTITAVTKD 24342
QY 241 SCV----- 244
Db 24343 SCVAVWKPASDGGAKIRNYLEKREKKONKWSVTTTEEIRETVFSVKNLTIEGLEYEFRV 24402
QY 245 ----- 244
Db 24403 KCENLGSESEWSEISETTPKSDVPIQAPHKFEELNLRVYQSNATLVCKVTHGPKPIV 24462
QY 245 ----- 244
Db 24463 KWRQKKEIIADGLKYRIQEPKGGYHQLIIASVTDODDATVYQVRATNOGGSVSGTASLEV 24522

```
QY 245 ----- 244
Db 24523 EVPAKIHLPKTLGMAVHALRGEVSIKIPFSGKDPVITWQGDLDINNGHYQVIVT 24582
QY 245 ----- 244
Db 24583 RSFTSLVFPNGVERKDAGFYVVCANKNRFQDKTVELDVADVPDPGRGVKVSVDVSRDSVN 24642
QY 245 ----- 244
Db 24643 LTWTEPASDGGSKITNYIIVEKCATTAERWLRVGOARETRYVINLFGKTSYQFRVIAENK 24702
QY 245 ----- 244
Db 24703 FGLSKPSEPSEPTITKEDKTRAMNYDEVDREVSMTKASHSTKELYEKYMAEDLGR 24762
QY 245 ----- 248
Db 24763 GEFIVHRCVETSSKTYMAKFVKVKGTDQVLVKEISILNIARHRNHLHSHESFESMEE 24822
QY 249 ----- 252
Db 24823 LVMIFEFISGLDIFERINTSAFELNBEREIVSVHVQVCEALQFLHSHNIGHFDIRPENIY 24882
QY 253 ----- 252
Db 24883 QTRRSSTIKIIFQARQLKPGDNFRLLFTAPEYAPEVHVDVSTATDMWSLGLTVYV 24942
QY 253 ----- 254
Db 24943 LLSGINFLAETNOQIENIMNAEYTFDEAFKEISTEAMDFVDRLLVKKRKSMTASEA 25002
QY 255 ----- 261
Db 25003 LQHPWLKQKIERVSTKVIKRLKRRYHTLKKDLNVMVSAARISCGGAIKRSQGVSAK 25062
QY 262 ----- 261
Db 25063 KVASIEIGPVSGIMHAVEGEGHKYCKIENYDQSTQVTWYFGVRLSENSEKEITY 25122
QY 262 ----- 261
Db 25123 EDGAILYVKDITKLDGTYRCKVWVNDYGEDSSYAELFKGVREVYDYCRRTMKKIKRR 25182
QY 262 ----- 261
Db 25183 TDTMRLLERPPEFTPLYNKTAIVGENVRGVITIVHPEPHVTWYKSGQIKPGDNDKKY 25242
QY 262 ----- 261
Db 25243 TFECDKGLYQLTINSVTDDAEVTVVARNKYGEDCKAKLTVTLHPPTDSTLRPMFKR 25302
QY 262 ----- 261
Db 25303 LLANAEQEQSGVCFEIRVSGIPPTLKWEKGQPLSLGNIEIIEGLDYALHIRTDL 25362
QY 262 ----- 261
Db 25363 PEDTGYRVATNTAGSTSCQAHQVERLRYKKQEFKSEERHVKQKIDKTLRMAEIL 25422
QY 262 ----- 261
Db 25423 SGTESVPLTQVAKEALREAAVLYKPAVSTKTVKGEFRLEIEEKKERKLMPYDVPEPRK 25482
QY 262 ----- 264
Db 25483 YKOTTIEDORIKQFVPMDSMKWYKKIRDQYEMPGK 25518
```


CC form heterodimers (SR-alpha:beta-1) that act as antagonist to the
CC cytokine by binding the cytokine to form a non-functional complex. The
CC receptor components are shared by cytokines such as the CNTF (ciliary
CC neurotrophic factor) family of cytokines. The invention provides the
CC basis for the development of IL-6 antagonists, as they show that if, in
CC the presence of a ligand, a non-functional intermediate complex,
CC consisting of the ligand, its alpha receptor and its beta-1 receptor
CC component, can be formed, it will effectively block the action of the
CC ligand. Effective antagonists of IL-6 or CNTF consist of heterodimers
CC of the extracellular domains of the alpha specificity determining
CC components of their receptors and the extracellular domain of gp130.
CC The resultant heterodimers, function as high-affinity traps, rendering
CC the cytokine inaccessible to form a signal transducing complex with the
CC native membrane-bound forms of their receptor. The nucleic acids and
CC polypeptides are useful for treating cytokine-related diseases or
CC disorders such as osteoporosis and primary and secondary effects of
CC cancer including multiple myeloma or cachexia.
XX
SQ Sequence 690 AA;
Query Match 94.0%; Score 1254; DB 21; Length 690;
Best Local Similarity 52.1%; Pred. No. 6.1e-36;
Matches 233; Conservative 4; Mismatches 8; Indels 202; Gaps 1;
QY 2 WLKAFY----- 7
DB 244 WNSFYRLRFELRYRAERSKFTTWMVKDLQHHCVHDAWSGLRHVVQLRAQEEFGQGW 303
QY 8 ----- 7
DB 304 SEWSPAMCTPWTSRPPAENEVSTPMQALTNKDDNILFRDSANATSLPVQAGAST 363
QY 8 ----- 7
DB 364 KGPSVFPLAPSKSTSGGTAALCLVKDYFPEPTVSWNSGALTSQVHTFPAVLQSSGLY 423
QY 8 -----DKVAEKLKEAFMDKTHCPAPPELLGGPSV 39
DB 424 SLSSVVTVPSSSLGTQTYICNVNHRKPSNTKDKKVEPKSCDKTHCPAPPELLGGPSV 483
QY 40 FLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSTY 99
DB 484 FLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSTY 543
QY 100 RVVSVTLVLHQDLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTK 159
DB 544 RVVSVTLVLHQDLNGKEYCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTK 603
QY 160 NOVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLDSDGFFLYSKLTVDKSRWQOG 219
DB 604 NOVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLDSDGFFLYSKLTVDKSRWQOG 663
QY 220 NVFSCVMHEALHNHYTQKSLSLSPCK 246
DB 664 NVFSCVMHEALHNHYTQKSLSLSPCK 690
RESULT 13
AAW70796
ID AAW70796 standard; protein; 859 AA.
XX
AC AAW70796;
XX
DT 03-FEB-1999 (first entry)
DE Human gp130-FC-His6 amino acid sequence.
XX gp130; cytokine antagonist; interleukin; gamma-interferon;
KW granulocyte macrophage colony-stimulating factor; J peptide;
KW transforming growth factor-beta.
XX
OS Synthetic.
OS Homo sapiens.

XX Key Location/Qualifiers
FH Protein 1..619
FT /note= "human gp130"
FT Misc-difference 2
FT /label= L2V
FT /note= "amino acid changed to accomodate a Kozak
FT sequence"
FT Peptide 1..22
FT /note= "signal peptide"
FT Misc-difference 620..621
FT /note= "Ser-Gly bridge"
FT Protein 662..853
FT /note= "from the Fc domain of human IgG1"
FT Disulfide-bond 632..635
FT Peptide 854..859
FT /note= "hexahistidine tag"
XX
PN US5844099-A.
XX
PD 01-DEC-1998.
XX
PF 27-NOV-1995; 95US-0563105.
XX
PR 27-NOV-1995; 95US-0563105.
PR 20-OCT-1993; 93US-0140222.
XX
PA (REGE-) REGENERON PHARM INC.
XX
PI Economides A, Stahl N, Yancopoulos GD;
XX
DR WPI; 1999-044669/04.
XX
PT Cytokine antagonists - comprising extracellular domains of
PT specificity-determining and signal-transducing components of
PT cytokine receptor
XX
PS Example 3; Fig 4; 46pp; English.
XX
CC The present sequence represents the amino acid sequence of human
CC gp130-FC-His6. The protein is used in the course of the invention. The
CC specification describes cytokine antagonists comprising only the
CC extracellular domain of the specificity-determining component of
CC the cytokine receptor and the extracellular domain of a
CC signal-transducing component of the cytokine receptor. The cytokine
CC is an interleukin (IL-1, IL-2, IL-3, IL-4, IL-5 or IL-15),
CC granulocyte macrophage colony-stimulating factor (GM-CSF),
CC gamma-interferon or transforming growth factor-beta (TGF-beta). The
CC antagonist is capable of binding the cytokine to form a nonfunctional
CC complex. The compounds have therapeutic activity as cytokine antagonists
CC and can also be used in assays for identifying novel agonists and
CC antagonists of cytokines.
SQ Sequence 859 AA;
Query Match 94.0%; Score 1254; DB 20; Length 859;
Best Local Similarity 60.5%; Pred. No. 1e-35;
Matches 233; Conservative 8; Mismatches 5; Indels 139; Gaps 3;
QY 1 DW-----LKAFF----- 6
DB 469 DWQEDGTVHRTYLRGNLAESKCYLITVPYADGPGSPESIRAYLKLQAPPSKGPVTRK 528
QY 7 -----YDKVAEKLKEAFM----- 19
DB 529 KVGKNEAVLEWDLQDPVDVQNGFIRNYTIFRTIIGNETAENVNDSSHTTEYTLSSLTSDTLY 588
QY 20 -----DKTHTCPPCPAPPELLGGPSVFL 41
DB 589 MYRMAAYTDEGGKDGPEFTTTPKFAQGEIESGEPKSCDKTHCPAPPELLGGPSVFL 648
QY 42 FPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSTYRV 101
|||||

```

XX AAW48650;
XX
XX 04-AUG-1998 (first entry)
XX
XX Heavy chain of hmb425 fused to TNF alpha.
XX
XX Antibody-cytokine fusion protein; tricistronic vector; chimeric;
XX TNF alpha; IL-2; IRES; internal ribosome entry site.
XX
XX Synthetic.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Region 1..494
XX /note= "Heavy chain of human mAb 425"
XX Region 495..652
XX /note= "TNF alpha"
XX
XX WO9811241-A1.
XX
XX 19-MAR-1998.
XX
XX 02-SEP-1997; 97WO-EP04765.
XX
XX 30-SEP-1996; 96EP-0115635.
XX
XX 16-SEP-1996; 96EP-0114920.
XX
XX (MERE ) MERCK PATENT GMBH.
XX
XX Bruenmer W, Burge C, Dunker R, Hauser H, Mielke C;
XX Rieke E, Von Hoegen I, Welge T;
XX
XX WPI: 1998-207400/18.
XX N-PSDB; AAV18096.
XX
XX Oligo:cistronic expression vector - useful for production of, e.g.
XX MAb425/TNF- $\alpha$  or MAb425/IL-2 antibody fusion protein
XX
XX Disclosure; Fig 15; 89pp; English.
XX
XX The present sequence represents a fusion protein comprising of TNF
XX alpha fused to the C-terminus of the heavy chain of the human
XX monoclonal antibody 425 (hmb425). The hmb425 has specificity for
XX the human EGF receptor. The invention claims for a new pMCDHAP
XX tricistronic vector (AAV18096) for the expression of an
XX antibody-cytokine fusion protein, hmb425-TNF alpha. The TNF alpha
XX sequence can be substituted by the IL-2 sequence. The vector also
XX contains a strong promoter/enhancer unit, a selection marker gene and at
XX least two poliovirus derived internal ribosomal entry site (IRES)
XX sequences. The vector can be expressed in mammalian host cells for the
XX production of heteromeric fusion proteins. This expression system is
XX claimed to produce the heteromeric proteins in high yields.
XX
XX Sequence 652 AA;
XX
XX Query Match 94.0%; Score 1254; DB 19; Length 652;
XX Best Local Similarity 54.4%; Pred. No. 5.3e-36;
XX Matches 234; Conservative 5; Mismatches 7; Indels 184; Gaps 2;
XX
XX 1 DWLKAF-----YD- 8
XX .:|:|
XX 65 EWIGEPNPNGRNYNEKFSKATMTVDSTNTAYMELSLRSEDVAVYCASRDYDYG 124
XX
XX 9 ----- 8
XX
XX 125 RYFDYWGQGLTVSSGEWILCAWALCPTPRSHGTTSLAATKGPVFPAPSSKSTSG 184
XX
XX 9 ----- 8
XX
XX 185 GTAAAGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVTVFSSSLGTQT 244
XX
XX 9 -----KVAEKLKAEAFMDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMISRTTP 56

```

```

Db 245 YICNVNHPKPSNTKVDKKVEPKSCDKTHCTCPAPELLGGPSVFLFPPKPKDTLMISRTTP 304
XX
QY 57 EYTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGK 116
XX
Db 305 EVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGK 364
XX
QY 117 EYKCKVSNKALPAPIEKTISKAKGQPREPOVYTTLPSSRDELTKNOVSLTCLVKGFYPSDI 176
XX
Db 365 EYKCKVSNKALPAPIEKTISKAKGQPREPOVYTTLPSSRDELTKNOVSLTCLVKGFYPSDI 424
XX
QY 177 AVESWGQPNENNYKTTTPVLDSDGSGFFLYSKLTVDKSRWQQGNFVSCSYMHEALHNHYT 236
XX
Db 425 AVESWGQPNENNYKTTTPVLDSDGSGFFLYSKLTVDKSRWQQGNFVSCSYMHEALHNHYT 484
XX
QY 237 QKSLSLSPGK 246
XX
Db 485 QKSLSLSPGK 494
XX
XX
XX RESULT 12
XX AAY92195
XX ID AAY92195 standard; protein; 690 AA.
XX
XX AC AAY92195;
XX
XX 01-AUG-2000 (first entry)
XX
XX Human IL-6R-alpha-C-gamma-1 fusion protein.
XX
XX IL-6R-alpha-C-gamma-1; cytokine; antagonist; CNTF; receptor;
XX fusion protein; cytostatic; immunomodulator; osteopathic.
XX
XX Synthetic.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Protein 1..358
XX /label= IL-6R-alpha
XX Peptide 359..360
XX /note= "Ala-Gly bridge"
XX Protein 361..690
XX /label= C-gamma-4
XX
XX WO200018932-A2.
XX
XX 06-APR-2000.
XX
XX 22-SEP-1999; 99WO-US22045.
XX
XX 25-SEP-1998; 98US-0101858.
XX 19-MAY-1999; 99US-0313942.
XX
XX (REGE-) REGENERON PHARM INC.
XX
XX Stahl N, Yancopoulos GD;
XX
XX WPI: 2000-293165/25.
XX
XX Isolated nucleic acid molecule for treating cytokine-related diseases
XX or disorders encodes a fusion polypeptide capable of binding a cytokine
XX to form a nonfunctional complex
XX
XX Example 4; Page -: 152pp; English.
XX
XX The invention concerns production of antagonists to any cytokine that
XX utilizes an alpha specificity determining component, which when combined
XX with the cytokine, binds to a first beta signal transducing component to
XX form a non-functional intermediate which then binds to a second beta
XX signal transducing component causing beta-receptor dimerization, the
XX soluble alpha specificity determining component of the receptor
XX (SR-alpha) and the extracellular domain of the first beta signal
XX transducing component of the cytokine receptor (beta-1) are combined to

```

```
PN WO200123431-A1.
XX 05-APR-2001.
XX 29-SEP-2000; 2000WO-JP06775.
XX 30-SEP-1999; 99JP-0278292.
XX (KYOW ) KYOWA HAKKO KOGYO KK.
XX Hanai N, Nakamura K, Niwa R;
XX WPI; 2001-266142/27.
XX Monoclonal antibodies against ganglioside GM2 combined with drugs,
XX radioisotopes or proteins for treatment and diagnosis of cancer -
XX Claim 43; Page 61-65; 80pp; Japanese.
XX The present invention relates to derivatives of an antibody against
XX ganglioside GM2. The antibody may be a monoclonal antibody or its
XX fragments. The antibody is combined with a radioactive isotope,
XX protein or small drug in the treatment and diagnosis of cancer.
XX Sequence 583 AA;
XX
Query Match 94.1%; Score 1255; DB 22; Length 583;
Best Local Similarity 56.2%; Pred. No. 3.8e-36;
Matches 234; Conservative 5; Mismatches 7; Indels 170; Gaps 2;
QY 1 DWLK-----AP 6
DB 35 DWVKQSQGGLEWGMGYIYPNNGGTGYNOKFKSKVTITVDTSTAYMELHSLRSEDYAVY 94
QY 7 Y-----7
DB 95 YCATYGHYGYWFAYWGGTLTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFP 154
QY 8 -----DKV 10
DB 155 EPVTVSNWNSGALTSGVHTFPAVLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHRKPSNTKV 214
QY 11 AEKLEAFMDKTHCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDP 70
DB 215 DKKVEPKSCDKTHCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDP 274
QY 71 EVKFNWYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVKSNKALPAP 130
DB 275 EVKFNWYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVKSNKALPAP 334
QY 131 IEKTSKAGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYPSDIAVWESNGOPENNY 190
DB 335 IEKTSKAGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYPSDIAVWESNGOPENNY 394
QY 191 KTTPLVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTQKLSLSLSPCK 246
DB 395 KTTPLVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTQKLSLSLSPCK 450
RESULT 10
AAG64473
ID AAG64473 standard; Protein; 473 AA.
XX
AC AAG64473;
XX
XX 25-SEP-2001 (first entry)
XX Human type antihuman IgE antibody H chain 3.
DE Human; antihuman IgE antibody; immunoglobulin; treating;
XX Human; antihuman IgE antibody; immunoglobulin; treating;
XX allergic disease.
XX Homo sapiens.
OS
```

```
XX Key Location/Qualifiers
FH Misc-difference 22
FT /note= "Encoded by CAG"
ET Misc-difference 129..141
FT /note= "Encoded by ccgtggggcc agggaaacacc ggtcgccttt
XX gactacgtc"
XX WO200151507-A1.
XX 19-JUL-2001.
XX 15-JAN-2001; 2001WO-JP00181.
XX 14-JAN-2000; 2000JP-0007061.
XX (SNOW ) SNOW BRAND MILK PROD CO LTD.
XX Washida N, Takahashi K, Satake T, Fujise N, Tanaka H, Kuriyama M;
XX WPI; 2001-442132/47.
XX N-PSDB; AAH47901.
XX New peptide used for screening human anti-human immunoglobulin E
XX monoclonal antibody useful in medical compositions for treating
XX allergies -
XX Example 6; Page 59-60; 70pp; Japanese.
XX The present sequence is that of a human type antihuman IgE antibody H
XX chain. The invention relates to a peptide useful in a method for
XX screening for human type antihuman IgE monoclonal antibodies
XX (AAH47897-AAH47904 encoding AAG64469-AAG64476) useful for preventing
XX and/or treating allergic disease.
XX Sequence 473 AA;
XX
Query Match 94.0%; Score 1254; DB 22; Length 473;
Best Local Similarity 55.5%; Pred. No. 2.5e-36;
Matches 233; Conservative 6; Mismatches 7; Indels 174; Gaps 2;
QY 1 DWLK-----4
DB 54 NWROVPQGGLEWMAIIPFGPPKYAOKFQGRVSLTADRSTNTAYMEMARLSRSDYAVY 113
QY 5 -----AFYD-----8
DB 114 YCARAPTLYDTEGTGPFYWGOGTPVAVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVK 173
QY 9 -----8
DB 174 DYFPEPVTVSNWNSGALTSGVHTFPAVLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHRKPS 233
QY 9 --KVAEKLKEAFMDKTHCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVDS 66
DB 234 NTKVDKVEPKSCDKTHCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVDS 293
QY 67 HEDPEVKFNWYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVKSNKA 126
DB 294 HEDPEVKFNWYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVKSNKA 353
QY 127 LPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYPSDIAVWESNGOP 186
DB 354 LPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYPSDIAVWESNGOP 413
QY 187 ENNYKTTPTPLVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTQKLSLSPCK 246
DB 414 ENNYKTTPTPLVLDSDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNYHTQKLSLSPCK 473
RESULT 11
AAW48650
ID AAW48650 standard; Protein; 652 AA.
```



```
Db 372 RWKSHLQNYTVNATKLTNLTNDRYLATLTVRNLVGKSDAAVLTIPACDFQATHPVMOLK 431
QY 5 AF-----6
Db 432 APFKDMLWVETTPRESVKYILEWCVLSDKAPCITDWOQEDGTVHRTYLRGNLAESKC 491
QY 7 -----YD-----
Db 492 YLITVTPVADGPGSPESIKAYLKQAPPSKGTPTVRTKVKNEAVLEWDLQDPVDVQNGFI 551
QY 9 -----8
Db 552 RNYTIFYRTIIGNETAVNVDSHSHTYTLSSLTSDTLVYMRMAAYTDEGKGDPFTTTP 611
QY 9 -----8
Db 612 KFAQGEIESGASTKGPSVFPPLAPSSKSTSGGTAALGCLVKDYFPEPTVSWNSGALTSKV 671
QY 9 -----KVAEKLKEAFMDKTHTCP 26
Db 672 HTFPAVLQSSGLYSLSSVTVPPSSSLGTQTYICNVNHNKPSNTKVDKVEPKSCDKTHTCP 731
QY 27 PCPAPELLGGPSVFLPPPKPDKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNA 86
Db 732 PCPAPELLGGPSVFLPPPKPDKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNA 791
QY 87 KTKPREQYNTYRVVSVTLVHQLDNLGKCKYKVSNNKALPAPTEKTIKAKGQPREPQ 146
Db 792 KTKPREQYNTYRVVSVTLVHQLDNLGKCKYKVSNNKALPAPTEKTIKAKGQPREPQ 851
QY 147 VYTLPPSRDELTKQVSLTCLVKGYPSDIAVWESNGQPNENYKTPPVLDSDGSFFLY 206
Db 852 VYTLPPSRDELTKQVSLTCLVKGYPSDIAVWESNGQPNENYKTPPVLDSDGSFFLY 911
QY 207 SKLTVDKSRWQGNVFCSSVHMEALHNNHYTQKSLSPGK 246
Db 912 SKLTVDKSRWQGNVFCSSVHMEALHNNHYTQKSLSPGK 951
```

```
RESULT 7
AA92186
ID AA92186 standard; protein: 951 AA.
XX
AC AA92186;
XX
DT 01-AUG-2000 (first entry)
XX
DE Human gp130-C-gamma-1.
XX
KW gp130-C-gamma-1; cytokine; antagonist; CNTF; receptor; fusion protein;
KW cytosolic; immunomodulator; osteopathic.
XX
OS Synthetic.
XX
OS Homo sapiens.
FH Key Location/Qualifiers
FT Protein 1..619
FT Peptide /label= gp130
FT Peptide 620..621
FT Protein /note= "Ser-Gly bridge"
FT Protein 622..951
XX
XX WO200018932-A2.
XX
XX
XX PD 06-APR-2000..
XX
XX PF 22-SEP-1999; 99WO-US22045.
XX
XX PR 25-SEP-1998; 98US-0101858.
XX
XX PR 19-MAY-1999; 99US-0313942.
XX
```

```
PA (REGE-) REGENERON PHARM INC.
XX
PI Stahl N, Yancopoulos GD;
DR
XX WPI; 2000-293165/25.
PT Isolated nucleic acid molecule for treating cytokine-related diseases
PT or disorders encodes a fusion polypeptide capable of binding a cytokine
PT to form a nonfunctional complex
XX
PS Example 4; Fig 9; 152pp; English.
XX
CC The invention concerns production of antagonists to any cytokine that
CC utilizes an alpha specificity determining component, which when combined
CC with the cytokine, binds to a first beta signal transducing component to
CC form a non-functional intermediate which then binds to a second beta
CC signal transducing component causing beta-receptor dimerization, the
CC soluble alpha specificity determining component of the receptor
CC (SR-alpha) and the extracellular domain of the first beta signal
CC transducing component of the cytokine receptor (beta-1) are combined to
CC form heterodimers (SR-alpha:beta-1) that act as antagonist to the
CC cytokine by binding the cytokine to form a non-functional complex.
CC Receptor components are shared by cytokines such as the CNTF (ciliary
CC neurotrophic factor) family of cytokines. The invention provides the
CC basis for the development of IL-6 antagonists, as they show that if, in
CC the presence of a ligand, a non-functional intermediate complex,
CC consisting of the ligand, its alpha receptor and its beta-1 receptor
CC component, can be formed, it will effectively block the action of the
CC ligand. Effective antagonists of IL-6 or CNTF consist of heterodimers
CC of the extracellular domains of the alpha specificity determining
CC components of the receptors and the extracellular domain of gp130.
CC The resultant heterodimers, function as high-affinity traps, rendering
CC the cytokine inaccessible to form a signal transducing complex with the
CC native membrane-bound forms of their receptor. The nucleic acids and
CC polypeptides are useful for treating cytokine-related diseases or
CC disorders such as osteoporosis and primary and secondary effects of
CC cancer including multiple myeloma or cachexia.
XX
SQ Sequence 951 AA;
```

```
Query Match 94.2%; Score 1257; DB 21; Length 951;
Best Local Similarity 37.0%; Pred. No. 1e-35; 5; Indels 394; Gaps 3;
Matches 237; Conservative 4; Mismatches 5;
QY 1 DW-----2
Db 312 DWSEASGITYEDRPSKAPSFWKIDPSHTQGYRTVQLVWKTLPPEANGKILDYEVTLT 371
QY 3 -----LK 4
Db 372 RWKSHLQNYTVNATKLTNLTNDRYLATLTVRNLVGKSDAAVLTIPACDFQATHPVMOLK 431
QY 5 AF-----6
Db 432 APFKDMLWVETTPRESVKYILEWCVLSDKAPCITDWOQEDGTVHRTYLRGNLAESKC 491
QY 7 -----YD-----8
Db 492 YLITVTPVADGPGSPESIKAYLKQAPPSKGTPTVRTKVKNEAVLEWDLQDPVDVQNGFI 551
QY 9 -----8
Db 552 RNYTIFYRTIIGNETAVNVDSHSHTYTLSSLTSDTLVYMRMAAYTDEGKGDPFTTTP 611
QY 9 -----8
Db 612 KFAQGEIESGASTKGPSVFPPLAPSSKSTSGGTAALGCLVKDYFPEPTVSWNSGALTSKV 671
QY 9 -----KVAEKLKEAFMDKTHTCP 26
Db 672 HTFPAVLQSSGLYSLSSVTVPPSSSLGTQTYICNVNHNKPSNTKVDKVEPKSCDKTHTCP 731
QY 27 PCPAPELLGGPSVFLPPPKPDKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNA 86
```

CC been constructed, which may be useful for treating neurological
CC disorders. The ephrin fusion proteins are preferably capable of binding
CC to Elk receptor and are especially Efl-6 antagonists. The fusion proteins
CC were constructed after it was demonstrated that similar improved
CC activities could be achieved using Tie-2 receptor ligands.
CC Angiopoietin-1 (Ang-1) is a Tie-2 receptor ligand and is an agonist for
CC Tie-2, whereas angiopoietin-2 (Ang-2) is a naturally occurring antagonist
CC of the Tie-2 receptor. The fibrinogen domains (FD) of Ang-1 and Ang-2 are
CC the receptor-binding domains and dimerized versions, of e.g. Ang-1-FD-Fc
CC (Ang-1 fibrinogen domain fused to an Fc domain), can bind to the Tie-2
CC receptor with much higher affinity than monomeric Ang-1-FD (dimerization
CC occurs between the Fc components of adjacent molecules). However,
CC Ang-1-FD-Fc is not able to induce phosphorylation (activate) the Tie-2
CC receptor on endothelial cells unless it is further clustered with goat
CC anti-human Fc antibodies. The novel fusion proteins, mutant versions of
CC Ang-1-FD and Ang-2-FD, were designed that were intrinsically more highly
CC clustered. Tie-2 agonist fusion proteins may be used as haematopoietic
CC factors. Tie-2 receptor antagonist fusion proteins may be used to
CC diagnose or treat, e.g. myeloproliferative or other proliferative
CC disorders of blood forming organs, e.g. thrombocythemias, polycythemias
CC and leukemias.

SQ Sequence 683 AA;

Query Match 94.2%; Score 1257; DB 21; Length 683;
Best Local Similarity 35.4%; Pred. No. 4 6e-36;
Matches 238; Conservative 3; Mismatches 4; Indels 427; Gaps 4;

QY 2 WLKA-----FYDKVAEL----- 14
|||
DB 12 WLVMVWALRLATPLAKNLEPVSWSLNPFLSGKGLVLYPKIGDKLDIICPRAEAGR 71
QY 15 ----- 14
DB 72 PYEYKLYLVPRQAAACSTVLDPNVLTNCNRPQEIRFTIKQEPSPNMYGLFEXKHHD 131
QY 15 -----KEA----- 17
|||
DB 132 YYITSTNGSLEGLNREGGVCRTMTKIMKVGQDPNAVTPQLTTSRPSREADNTVMK 191
QY 18 ----- 17
DB 192 ATQAPGSGSLGSDGKHETVNOEKGSGPGASGSGDPDGFNKGPKGNLEPVSWSL 251
QY 18 ----- 17
DB 252 NPKFLSGGLVLYPKIGDKLDIICPRAEAGRPEYKLYLVPRQAAACSTVLDPNVLT 311
QY 18 -----FM----- 19
|||
DB 312 CNRPEQIRFTIKQEPSPNMYGLFEXKHHDYYITSTNGSLEGLNREGGVCRTMTKI 371
QY 20 ----- 19
DB 372 IMKVGQDPNAVTPQLTTSRPSKEADNTVMKATQAPGSRGSLGSDGKHETVNOEKGSGP 431
QY 20 -----DKTHTCPCAPPELLGGPSVFLFPKPKDTLMISR 54
DB 432 GASGSGSDPDGFNKGPKGPKSCDKTHTCPCAPPELLGGPSVFLFPKPKDTLMISR 491
QY 55 TPEVTCVVVDVSHEDPEVKFNKYVDGVEVHNATKPREEQYNSTYRVSVLTVLHQDWLN 114
DB 492 TPEVTCVVVDVSHEDPEVKFNKYVDGVEVHNATKPREEQYNSTYRVSVLTVLHQDWLN 551
QY 115 GKEYKCKVNKALPAPIETKISKAKQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPS 174
DB 552 GKEYKCKVNKALPAPIETKISKAKQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPS 611
QY 175 DTAVESNGQPNENYKTPPVLDSDGSPFLYSKLTVDKSRWQOQNVFSCSVNHEALHNH 234
DB 612 DTAVESNGQPNENYKTPPVLDSDGSPFLYSKLTVDKSRWQOQNVFSCSVNHEALHNH 671
QY 235 YTKSLSLSPGK 246

DB 672 YTKSLSLSPGK 683
|||||||
RESULT 6
AAW70798
ID AAW70798 standard; protein; 951 AA.
XX
AC AAW70798;
XX
DT 03-FEB-1999 (first entry)
XX
DE Human gp130-C-gamma-1 amino acid sequence.
XX
KW gp130; cytokine antagonist; interleukin; gamma-interferon;
KW granulocyte macrophage colony-stimulating factor; J peptide;
KW transforming growth factor-beta.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Protein 1..619
FT /note= "human gp130"
FT Protein 662..651
FT /note= "from the constant region of human IgG1"
XX
PN US5844099-A.
XX
PD 01-DEC-1998.
XX
PF 27-NOV-1995; 95US-0563105.
XX
PR 27-NOV-1995; 95US-0563105.
PR 20-OCT-1993; 93US-0140222.
XX
PA (REGE-) REGENERON PHARM INC.
XX
PI Economides A, Stahl N, Yancopoulos GD;
XX
WPI; 1999-044669/04.
XX
PT Cytokine antagonists - comprising extracellular domains of
PT specificity-determining and signal-transducing components of
PT cytokine receptor
XX
PS Example 4; Fig 9A-B; 46pp; English.
XX
CC The present sequence represents the amino acid sequence of human
CC gp130-C-gamma-1. The protein is used in the course of the invention.
CC The specification describes cytokine antagonists comprising only the
CC extracellular domain of the specificity-determining component of
CC the cytokine receptor and the extracellular domain of a
CC signal-transducing component of the cytokine receptor. The cytokine
CC is an interleukin (IL-1, IL-2, IL-3, IL-4, IL-5 or IL-15),
CC granulocyte macrophage colony-stimulating factor (GM-CSF),
CC gamma-interferon or transforming growth factor-beta (TGF-beta). The
CC antagonist is capable of binding the cytokine to form a nonfunctional
CC complex. The compounds have therapeutic activity as cytokine antagonists
CC and can also be used in assays for identifying novel agonists and
CC antagonists of cytokines.
SQ Sequence 951 AA;

Query Match 94.2%; Score 1257; DB 20; Length 951;
Best Local Similarity 37.0%; Pred. No. 1e-35;
Matches 237; Conservative 4; Mismatches 5; Indels 394; Gaps 3;
QY 1 DW----- 2
|||
DB 312 DWSEASGTYEDRPSKAPSFYKIDPSHTQGYRTVOLVWKLPPFEANGKILDYEVTLT 371
QY 3 -----LK 4

XX Protein sequence of primatised form of the heavy chain of 16C10 antibody.
DE Human; macaque monkey; light chain; primatised antibody; 16C10 antibody;
KW neuroprotective; apoptosis inducer; allergy; CD28 receptor antagonist;
KW B7.1 antigen; CD80; B7.2 antigen; CD86; B cell cancer; metastasis;
KW tumour; B cell lymphoma; B cell leukaemia; autoimmune disease;
KW graft-vs-host disease; immunosuppression; organ rejection;
KW interleukin-2; IL-2; mutant; mutein.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Macaca Sp.
OS Synthetic.
XX WO200189567-A1.
XX 29-NOV-2001.
XX
XX 22-MAY-2001; 2001WO-US16364.
XX
XX 22-MAY-2000; 2000US-0576424.
XX (IDEC-) IDEC PHARM CORP.
XX
XX Anderson DR, Hanna N, Brams P;
XX
XX WPI; 2002-089895/12.
XX N-PSDB; AAS17247.
XX
XX Use of monoclonal antibody which specifically binds to B7.1 antigen
PT CD80 and/or B7.2 antigen CD86 for inducing apoptosis of B7+ cells,
PT treating cancer, graft-vs-host disease and autoimmune disease as
PT allergy
XX
XX Example 8; Fig 5b; 89pp; English.
XX
XX The present invention relates to a new use of a monoclonal antibody
CC which specifically binds to B7.1 antigen (CD80) and/or B7.2 antigen
CC (CD86) for inducing the apoptosis of B7+ cells. The invention is
CC useful for treating diseases such as B cell cancer, lymphoma, a
CC cancer where B cells promote the growth and/or metastasis of tumours,
CC B cell lymphoma, B cell leukaemia, and autoimmune diseases such as
CC idiopathic thrombocytopenia purpura, systemic lupus, erythematosis,
CC type 1 diabetes mellitus, rheumatoid arthritis, psoriasis, aplastic
CC anaemia, inflammatory bile disease, allergy, multiple sclerosis
CC or graft-vs-host disease. The antibody is useful for immunosuppression
CC in a human or animal and for treating or preventing resistance to or
CC rejection of transplanted organ or tissue for treating proliferative
CC and hyperproliferative diseases, for treating reversible obstructive
CC airways disease, intestinal inflammations and allergies e.g. Crohn's
CC disease and ulcerative colitis, food-related allergies e.g. migraine,
CC rhinitis and eczema, and other types of allergies. The present protein
CC sequence represents the heavy chain of 16C10, a primatised antibody
CC used in the invention to induce apoptosis and inhibit production of
CC interleukin-2 (IL-2).
XX
XX Sequence 476 AA;
XX
XX Query Match 94.3%; Score 1258; DB 23; Length 476;
XX Best Local Similarity 56.7%; Pred. No. 1.8e-36;
XX Matches 233; Conservative 5; Mismatches 8; Indels 165; Gaps 1;
XX
XX 1 DMLKAFYD----- 8
XX :||:
DB 66 EWIGSFYSSGNTYNNPSLKSQVTISTDTKNQFSLKNSMTAATVYICVRDLFSV 125
XX 9 ----- 8
DB 126 GMVYNNWFDWGPGLVTVSSASTKGPVFLAPSSKSTSGTAALGLVLDYFPEPVTV 185
XX 9 -----KVAELK 15
DB 186 SWNSGALTSGVHTFFPAVLQSLGSLSVTVVPSSSLGTQTYICNVNHPKSNKTKYDKRAE 245

QY 16 EAFMDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFN 75
DB PKSCDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFN 305
QY 76 WYDGVVEVHNAKTPREQYINSTRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTI 135
DB WYDGVVEVHNAKTPREQYINSTRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTI 365
QY 136 SKAKGQPREPOVYITLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPP 195
DB SKAKGQPREPOVYITLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPP 425
QY 196 VLDSDGSFELYSKLTVDKSRWQOGNMFVSCSVNHEALHNHYTQKSLSLSPGK 246
DB VLDSDGSFELYSKLTVDKSRWQOGNMFVSCSVNHEALHNHYTQKSLSLSPGK 476

XX RESULT 5
XX AAY96781
XX ID AAY96781 standard; Protein; 683 AA.
XX AC AAY96781;
XX DT 26-SEP-2000 (first entry)
XX XX Ephrin-B1-Ephrin-B1-FC fusion protein.
DE Ephrin-B1; ELK receptor; ligand; dimer; FC domain; fusion protein;
KW Ephrin-B1; ELK receptor; ligand; dimer; FC domain; fusion protein;
KW Efl-6 antagonist; neurological.
XX Chimeric - Homo sapiens.
OS Chimeric - Synthetic.
XX
XX Key Location/Qualifiers
XX Peptide 1..29
XX /label= Signal_peptide
XX Protein 30..237
XX /note= "derived from ELK-L Ectodomain 1"
XX /label= ELK-L_ectodomain_1
XX Peptide 238..240
XX /label= Bridging_peptide
XX Protein 241..448
XX /label= ELK-L_Ectodomain_2
XX Peptide 449..451
XX /label= Bridging_peptide
XX Region 452..683
XX /label= Human_IgG1_Fc_region
XX WO200037642-A1.
XX 29-JUN-2000.
XX 23-DEC-1999; 99WO-US30900.
XX 23-DEC-1998; 98US-0113387.
XX (REG-) REGENERON PHARM INC.
XX Davis SJ, Gale NW, Yancopoulos GD, Stahl N;
XX WPI; 2000-442670/38.
XX N-PSDB; AAA51345.
XX Polynucleotide encoding a fusion polypeptide, useful for promoting
PT differential function and influencing phenotype, comprises two subunits
PT containing at least one copy of the receptor binding domain of a ligand
XX Example 12; Fig 14A-E; 97pp; English.
XX
XX Production of homogenous forms of clustered ligands is broadly applicable
CC to improve the affinity and/or increase the activity of a ligand as
CC compared to the native form of the ligand. Ephrin fusion proteins have

CC B7.1 antigen monoclonal antibody 16C10. Cloned 16C10 light and heavy variable genes (see also AAT62512 and AAT62513) are inserted into an expression vector (pref. NEOSPLA) which contains human light and heavy chain constant region genes to allow prodn. of the primatised antibody in e.g. CHO cells. Primatised 7C10 and 7B6 anti-B7.1 antibodies have also been produced. (see also AAW01817-20). The primatised antibodies inhibit the B7:CD28 pathway, making them useful immunosuppressants for the treatment of autoimmune disorders and graft-versus-host disease.

XX CC
XX SQ Sequence 476 AA;

Query Match 94.3%; Score 1258; DB 18; Length 476;
Best Local Similarity 56.7%; Pred. No. 1.8e-36;
Matches 233; Conservative 5; Mismatches 8; Indels 165; Gaps 1;

QY 1 DWLKAFYD----- 8
: : : : :
Db 66 EWIGFYSSSGNTYYNPSLKSVTISTDTSKNQFSLKLSMTAADTAVYYCVRDLFSVY 125
: : : : :
QY 9 ----- 8
: : : : :
Db 126 GMVYNNWFDVWGPGVLVTSSASTKGPSVFPLAPSSKTSGGTAALGCLVKDYFPEPTV 185
: : : : :
QY 9 -----KVAEKLK 15
: : : : :
Db 186 SNNSGALTSGVHTFPAVLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHHKPSNTKVDKKA 245
: : : : :
QY 16 EAFMDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFN 75
: : : : :
Db 246 PKSCDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFN 305
: : : : :
QY 76 WYVDGVEVHNATKPREEQYNSTYRVSVSLTVLHODWLNKGEYCKKVSNNKALPAPIEKTI 135
: : : : :
Db 306 WYVDGVEVHNATKPREEQYNSTYRVSVSLTVLHODWLNKGEYCKKVSNNKALPAPIEKTI 365
: : : : :
QY 136 SKAKGPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEVESNGQPNNTKTPP 195
: : : : :
Db 366 SKAKGPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEVESNGQPNNTKTPP 425
: : : : :
QY 196 VLDSGFSFLYSLKLTVDKSRWQGNVFCVSMHEALHNNHYTKSLSPGK 246
: : : : :
Db 426 VLDSGFSFLYSLKLTVDKSRWQGNVFCVSMHEALHNNHYTKSLSPGK 476
: : : : :
RESULT 3
AAW63765
ID AAW63765 standard; Protein; 476 AA.
XX AC AAW63765;
XX DT 29-SEP-1998 (first entry)
XX DE Macaque primatized 16C10 heavy chain protein.
XX KW Monoclonal antibody; Mab; macaque; heavy chain; primate; antigen; CD80;
KW CD86; inhibitor; immunosuppressant; treatment; autoimmune disease; IL-2;
KW T cell/B cell interaction; tumour; inflammation; imaging agent; vaccine;
KW immunogen; anti-idiotype reagent; interleukin-2; IgG; immunoglobulin G;
XX T cell proliferation.
OS Macaca fascicularis.
XX KW W09819706-A1.
XX PN 14-MAY-1998.
XX PD 29-OCT-1997; 97WO-US19906.
XX PF 08-NOV-1996; 96US-0746361.
XX PR (IDEC-) IDEC PHARM CORP.
XX PA
XX DT

PI Anderson DR, Brans P, Hanna N;
XX WPI: 1998-286601/25.
DR N-PSDB; AAV35489.
XX New monoclonal antibodies specific for B7.1 or B7.2 antigens and inhibiting binding to CD28 - useful as specific immunosuppressants for treating diseases that involve interactions between T and B cells, e.g. graft rejection or tumours

PT
PT
XX
PS Example 7; Fig 5b; 87pp; English.

XX This sequence represents a primatized form of the antibody 16C10 heavy chain from macaque. This sequence is used in a method which studies new monoclonal antibodies (Mab's) that bind selectively to B7.1 (CD80) or to B7.2 (CD86) antigens and inhibits binding of these antigens to CD28. Such Mab's are specific immunosuppressants for treatment of diseases involving T cell/B cell interactions, particularly autoimmune disease, specifically idiopathic thrombocytopenia purpura, systemic lupus erythematosus, type I diabetes mellitus, rheumatoid arthritis, psoriasis, aplastic anaemia, inflammatory bowel disease, allergy and multiple sclerosis, graft vs. host diseases, B cell lymphoma, infections (including by human immune deficiency virus) or inflammatory disease and tumours. Optionally the Mab can be conjugated to a drug or toxin. Mab's, or their fragments, can also be used as imaging agents and as vaccines or immunogens to develop anti-idiotype reagents. Mab's are optionally combined with other proteins or small molecule immunosuppressants. Blocking B7/CD28 interactions induces long-term, antigen-specific immunosuppression, i.e. it inhibits production of interleukin-2 (IL-2), T cell proliferation and antigen-specific immunoglobulin G (IgG) responses.

XX SQ Sequence 476 AA;

Query Match 94.3%; Score 1258; DB 19; Length 476;
Best Local Similarity 56.7%; Pred. No. 1.8e-36;
Matches 233; Conservative 5; Mismatches 8; Indels 165; Gaps 1;

QY 1 DWLKAFYD----- 8
: : : : :
Db 66 EWIGFYSSSGNTYYNPSLKSVTISTDTSKNQFSLKLSMTAADTAVYYCVRDLFSVY 125
: : : : :
QY 9 ----- 8
: : : : :
Db 126 GMVYNNWFDVWGPGVLVTSSASTKGPSVFPLAPSSKTSGGTAALGCLVKDYFPEPTV 185
: : : : :
QY 9 -----KVAEKLK 15
: : : : :
Db 186 SNNSGALTSGVHTFPAVLQSSGLYSLSSVTVTPSSSLGTQTYICNVNHHKPSNTKVDKKA 245
: : : : :
QY 16 EAFMDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFN 75
: : : : :
Db 246 PKSCDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFN 305
: : : : :
QY 76 WYVDGVEVHNATKPREEQYNSTYRVSVSLTVLHODWLNKGEYCKKVSNNKALPAPIEKTI 135
: : : : :
Db 306 WYVDGVEVHNATKPREEQYNSTYRVSVSLTVLHODWLNKGEYCKKVSNNKALPAPIEKTI 365
: : : : :
QY 136 SKAKGPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEVESNGQPNNTKTPP 195
: : : : :
Db 366 SKAKGPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEVESNGQPNNTKTPP 425
: : : : :
QY 196 VLDSGFSFLYSLKLTVDKSRWQGNVFCVSMHEALHNNHYTKSLSPGK 246
: : : : :
Db 426 VLDSGFSFLYSLKLTVDKSRWQGNVFCVSMHEALHNNHYTKSLSPGK 476
: : : : :
RESULT 4
AAU11646
ID AAU11646 standard; Protein; 476 AA.
XX AC AAU11646;
XX DT 12-MAR-2002 (first entry)

PI Kainoh M, Tanaka T;
 XX WPI; 1998-427881/36.
 DR N-PSDB; AAV33772.
 XX Integrin-immunoglobulin chimeric protein heterodimer complexes as
 PT platelet substitutes - contain the alpha and beta integrin chains
 PT associated in stable state and bind to extracellular matrix in the
 PT presence of plasma components
 XX
 PS Claim 7; Pages 42-50; 87pp; Japanese.
 XX This represents an integrin alpha-4 chain. The invention provides
 CC integrin-immunoglobulin chimeric protein heterodimer complexes that
 CC comprise an integrin alpha or beta chain associated with an
 CC immunoglobulin light or heavy chain. These chimeric proteins form
 CC heterodimer complexes, in particular with a chimeric protein containing
 CC an integrin alpha chain and an immunoglobulin chain with a chimeric
 CC protein containing an integrin beta chain and an immunoglobulin chain;
 CC the immunoglobulin chain in each case may be a heavy chain, or one of the
 CC two may be a light chain. The integrin alpha chain is preferably alpha 4
 CC or alpha 2 and the integrin beta chain is preferably beta 1. Animal cells
 CC transformed with vectors containing the DNA coding for the above chimeric
 CC proteins can be used in the preparation of the chimeric proteins and
 CC their heterodimer complexes. The heterodimer complexes, which are useful
 CC for testing potential promoters and inhibitors of the binding of
 CC integrins to their ligands, function as blood platelet substitutes and
 CC hemostatics and as diagnostic agents.
 XX
 SQ Sequence 1218 AA;

Query Match 94.5%; Score 1261; DB 19; Length 1218;
 Best Local Similarity 24.2%; Pred. No. 1.3e-35;
 Matches 236; Conservative 4; Mismatches 3; Indels 734; Gaps 3;

QY 4 KAFYDK-----VAE 12
 |||||
 DB 242 KAFDKQNVKFGSYLGISVGAGHFRSQTTEVVGAPQHEQIGKAYIFSIDEKELNHL 301
 QY 13 KIK----- 15
 :||
 DB 302 EMKGKLGSYFGASCAVDLNADGFSLLVGAPMQSTIREGRVFYINSGGAYVNAME 361
 QY 16 ----- 15
 DB 362 TNLVGSCKYARFGESIVNLGDIIDNGFEDVAIGAPQEDDLQGAIIYINGRADGISSTFS 421
 QY 16 ----- 15
 DB 422 QRIEGLQISKLSLMSGQISQIDADNNGYVDVAVGAFRSDSAVLLRTPVIVDASLSH 481
 QY 16 ----- 15
 DB 482 PESVNRKFDVCVNGWSPVCIDTLCSYKKGKVPYIIVLFYNMSLDVNRKAESPFRYF 541
 QY 16 -----EAF----- 19
 :|||
 DB 542 SSGNTSDVITGSIQVSSREANCRTHQAFMRKDVDRILTPIQTEAAYHLGPHVSKRSTEE 601
 QY 20 ----- 19
 DB 602 FPPLOPILQOKKEKIMKKTINFARCAHENCADLQVSAKIGFLKPHENKTYLAVGSMK 661
 QY 20 ----- 19
 DB 662 TLMLNVLNAGDAYETTLHVKLPVGLYFIKILEEEKQINCEVTDNSGVVQDLSIGY 721
 QY 20 ----- 19
 DB 722 IYVDHLRIDISFLLDVSSLSRAEEDLSITVHATCENEEMDNLKHRSRTVAIPLKVEVK 781
 QY 20 ----- 19

DB 782 LTVHGFVNPTSFVYGSNDENEPETCMVEKMNLTFFHVINTGNSMAPNVSVIWPNSFSQP 841
 QY 20 ----- 19
 DB 842 TDKLENILOVQTTGCEHFENYQVRCALAEQQKSAQMTLKGIVRFLSKDKRLLLYCIKADP 901
 QY 20 ----- 19
 DB 902 HCLFNLNCGKMGESKEASVHIQLEGRPSILEMDETSALKFEIRATGPRPEPNRVIELNK 961
 QY 20 -----DKTHTCPCPAPELLGGPSVFLFPKPKDT 49
 |||||
 DB 962 DENVAHVLEGLHQRKRYFTDPEEPKSCDKTHTCPCPAPELLGGPSVFLFPKPKDT 1021
 QY 50 LMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSTYRVVSVLTVLH 109
 |||||
 DB 1022 LMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKPREQYNSTYRVVSVLTVLH 1081
 QY 110 QDWLNGKEYKCKVSKNKAAPLPIETKTISKAKGQPREPQVYVTLPPSRDELTKNOVSLTCLVK 169
 |||||
 DB 1082 QDWLNGKEYKCKVSKNKAAPLPIETKTISKAKGQPREPQVYVTLPPSRDELTKNOVSLTCLVK 1141
 QY 170 GFYPDSIAEVESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFCVMHE 229
 |||||
 DB 1142 GFYPDSIAEVESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFCVMHE 1201
 QY 230 ALHNHYTQKSLSLSPGK 246
 |||||
 DB 1202 ALHNHYTQKSLSLSPGK 1218

RESULT 2
 AAW01822
 ID AAW01822 standard; Protein; 476 AA.
 XX
 AC AAW01822;
 XX
 DT 25-MAY-1997 (first entry)
 XX
 DE Primatised anti-human B7.1 antigen antibody 16C10 heavy chain.
 XX
 KW Monoclonal antibody; cynomolgus monkey; macaque; 16C10;
 KW primatised antibody; B7 antigen; CD28; immunosuppressive;
 KW autoimmune disease; idiopathic thrombocytopaenia purpura;
 KW systemic lupus erythematosus; rheumatoid arthritis; psoriasis;
 KW type 1 diabetes mellitus; graft versus host disease;
 KW hetero-hybridoma; transfectoma.
 XX
 OS Chimeric Macaca cynomolgus;
 OS Chimeric Homo sapiens.
 XX
 PN WO9640878-A1.
 XX
 PD 19-DEC-1996.
 XX
 PF 06-JUN-1996; 96WO-US10053.
 XX
 PR 07-JUN-1995; 95US-0487550.
 XX
 PA (IDEC-) IDEC PHARM CORP.
 XX
 PI Anderson DR, Brams P, Hanna N, Shestowsky WS;
 XX
 DR WPI; 1997-108638/10.
 DR N-PSDB; AAT62513.
 XX
 PT Monkey monoclonal antibody binding human B7.1 or B7.2 antigen -
 PT useful for treating autoimmune disease or graft-versus-host disease
 XX
 PS Claim 14; Fig 10B; 81pp; English.
 XX
 CC 2 Polypeptides (AAW01821 and AAW01822) respectively comprise primatised
 CC forms of the light and heavy chains of cynomolgus monkey anti-human

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:21 ; Search time 28.9412 Seconds
(without alignments)
1132.630 Million cell updates/sec

Title: 7LINK2

Perfect score: 1334

Sequence: 1 DMLKAFYDKVAEKLKEAFMD.....MHEALHNYTKSLSLSPGK 246

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

1:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1261	94.5	1218	19	AAW70539
2	1258	94.3	476	18	AAW01822
3	1258	94.3	476	19	AAW63765
4	1258	94.3	476	23	AAU11646
5	1257	94.2	683	21	AAAY96781
6	1257	94.2	951	20	AAW70798
7	1257	94.2	951	21	AAAY97186
8	1257	94.2	961	21	AAAY92187
9	1255	94.1	583	22	AAAB83156
10	1254	94.0	473	22	AAAG64473

11	1254	94.0	652	19	AAW48650	Heavy chain of hma
12	1254	94.0	690	21	AAAY92195	Human IL-6R-alpha-
13	1254	94.0	859	20	AAW70796	Human gp130-Fc-His
14	1254	94.0	859	21	AAAY92184	Human gp130-Fc-His
15	1254	94.0	1158	21	AAAY92205	Fusion polypeptide
16	1254	94.0	1168	21	AAAY92204	Fusion polypeptide
17	1253	93.9	471	21	AAAY5030	HUMAN OCR10-Fc fus
18	1253	93.9	475	13	AAAR20057	Heavy chain of 3D6
19	1253	93.9	475	18	AAAL1639	Human anti-RSV mon
20	1253	93.9	475	22	AAAG63640	Human anti-RSV mon
21	1253	93.9	567	23	AAEL13733	Amino acid sequenc
22	1253	93.9	754	21	AAAB11691	Human zaiphalr/1g
23	1253	93.9	787	21	AAAB11693	Human secreted K10
24	1253	93.9	1367	19	AAW70542	Integrin alpha-2 c
25	1252	93.9	449	14	AAAR43339	Completely humanis
26	1252	93.9	449	19	AAW49816	Amino acid sequenc
27	1252	93.9	475	17	AAAR93553	Monoclonal antibod
28	1252	93.9	595	20	AAW86003	Anti-574 single ch
29	1251	93.8	462	21	AAAB29408	Human monoclonal a
30	1251	93.8	470	21	AAAB08026	A dimeric anti-CD2
31	1251	93.8	473	22	AAAG64475	Human type antihum
32	1251	93.8	475	18	AAW11641	Human anti-RSV mon
33	1251	93.8	476	20	AAW88464	Monoclonal antibod
34	1250	93.7	399	21	AAAY70867	Human interferon-b
35	1250	93.7	423	21	AAAY70869	Human interferon-b
36	1250	93.7	446	15	AAAS58753	VCAM 2D-IgG. Homo
37	1250	93.7	446	20	AAAY23986	VCAM 2D-IgG, a sol
38	1250	93.7	446	20	AAAY01037	VCAM 2D-IgG protei
39	1250	93.7	446	20	AAW96743	A VCAM 2D-IgG1 fus
40	1250	93.7	473	22	AAAG64469	Human type antihum
41	1250	93.7	473	22	AAAG64471	Human type antihum
42	1250	93.7	476	14	AAAR31023	Antibody D heavy c
43	1250	93.7	481	13	AAAR24442	Sequence of antibo
44	1250	93.7	592	22	AAAB83838	Amino acid sequenc
45	1250	93.7	963	19	AAW70540	Integrin beta-1 ch

ALIGNMENTS

RESULT 1

AAW70539
ID AAW70539 standard; Protein; 1218 AA.

XX AC AAW70539;

XX DT 26-JAN-1999 (first entry)

XX DE Integrin alpha-4 chain.

XX KW Integrin; alpha-4 chain; immunoglobulin; chimeric; heterodimer complex;

XX KW inhibitor; binding; ligand; blood platelet; hemostatic; diagnostic agent; human.

XX OS Homo sapiens.

XX FH Key

XX FT Peptide Location/Qualifiers

XX FT Protein 1...39 /note= "signal peptide"

XX FT Protein 40..1218 /note= "mature protein"

XX PN W09832771-A1.

XX PD 30-JUL-1998.

XX PF 29-JAN-1998; 98WO-JP00370.

XX PR 29-AUG-1997; 97JP-0234544.

XX PR 29-JAN-1997; 97JP-0015118.

XX PR (TORA) TORAY IND INC.

XX PA


```
Qy 261 ----- 260
Db 23978 SAKKSNKERIKDKQFTIGGLEATEYEFVFAENETGLSRPRRTAMSIKTLTSGEAPG 24037
Qy 261 ----- 260
Db 24038 IRKEMKDVTTKLGEAQLSCQIVGRPLDIKWYRFGKELIQSRKYKMSDGRTHLTVM 24097
Qy 261 ----- 260
Db 24098 BEQDEGVYTCIATNEVEGETSSKLLQATPQFHPGYPLEKEYYGAVGSTLRHLHVMYIG 24157
Qy 261 ----- 260
Db 24158 RVPVMTWPHGQKLLONSENITTENTHYTHLVKMNQVRKTHAGYKVQLSNVFGIVDAI 24217
Qy 261 ----- 260
Db 24218 LDVEIQDKPKDPTGPIVIEALLKNSAVISWKPADDGGSWITNVVVEKCEKEGAEWQLV 24277
Qy 261 ----- 260
Db 24278 SSAISVYTCRIVNLNENAGYFRVSAQNTFGISDPLEVSSVWIIKSPFEKPGAPGPTIT 24337
Qy 261 ----- 260
Db 24338 AVTKDCVWAWKPPASDGGAKIRNYVLEKREKKQNKWISVTTEEIRETVFSVKNLIEGLE 24397*
Qy 261 ----- 260
Db 24398 YEFVKCENLGSESEISEIPTPKSDVP IQAPHEKEELRNLRVRYQSNATLVCKVTGH 24457
Qy 261 ----- 260
Db 24458 PKPIVWYRQKEIIADGLKYRQEFKGYHQLIIASVTDATVYQVRATNOGGSVSGT 24517
Qy 261 ----- 260
Db 24518 ASLEVEVPAKIHLPKLTLEGVAVHALRGEVVSIKIPFSGKPDVITWQKQDLDNNGHY 24577
Qy 261 ----- 260
Db 24578 QVIVTRSFSLVPNGVERKDAGFYVVCARNRFGIDQKTVELDADVDPDPPRGVKVSDVS 24637
Qy 261 ----- 260
Db 24638 RDSVNLWTPEPASDGGSKITNYIVEKCATAERLVRGQARETRYVINLFGKTSYQFRV 24697
Qy 261 ----- 260
Db 24698 IAENKFGLSKPSPTITTKEDKTRAMNYDEEVDETREVSMTKASHSSTKELYEKMIA 24757
Qy 261 ----- 260
Db 24758 EDLGRGEFIVHRCVETSSKTTYNAKFVKVGTQDVLVKKKISILNIARHRNILLHESF 24817
Qy 261 ----- 260
Db 24818 ESMEELVMIFEFGSLDIFERINTSAFELNEREIVSVHQVCEALQFLHSHNIGHFDIRP 24877
Qy 261 ----- 260
Db 24878 ENIIYQTRRSSTIKIIEFGAQLKPGDNFRLLFTAPEYYAPEVHQHVDVSTATDMWSLG 24937
Qy 261 ----- KEAF 264
:
Db 24938 TLVYVLLSGINPFLAETNQOIENIMNAEYTFDEEAF 24974
```

Search completed: April 21, 2003, 10:52:25
Job time : 104.824 secs

Db 21758 GPPTVVKVTDTSKTTVSLWSKPVFDGGMEIIGYIEMCKADLDGWHKVNACVKTRYT 21817
QY 248 ----- 247
Db 21818 VTDLQAGEYKFRVSAINGACKGDSCEVTGTIKAVDRUTAPELDIDANFKQTHVVVRAGAS 21877
QY 248 ----- 247
Db 21878 IRLFIAYOGRPTPTAVWSKPSNLSLRADIIHTDTSFSLTVENCNRNDAGKYTLTVENNS 21937
QY 248 ----- 247
Db 21938 GSKSITFTVKVLDTPGPPGPIITFKDVTGRGSATLMDAPLLDGGARIHHYVVEKREASRRS 21997
QY 248 ----- 247
Db 21998 QWVISEKCTROIFKVNOLAEGVPYIFRVSAVNEYGVPYEMPEPIVATEQPAPPRLDV 22057
QY 248 ----- 247
Db 22058 VDTSSSAVLAWLKPDDHGGSRITCYLLEMRQKGSDFWVEAGHTKQLFTTVERLVEKTEY 22117
QY 248 ----- 247
Db 22118 EFRVAKNDAGYSEPREAFSSVIIKEPOIPTADLTGITNQLITCKAGSPFTIDVPISGR 22177
QY 248 ----- 247
Db 22178 PAPKVTWKEMLUKETDRVSITTTKDRITLTIVKDSMRGDSGRYFLTLENTAGVKTFEVS 22237
QY 248 ----- 247
Db 22238 VVIGRCPVTCPIEVSSVSAESCVLSWGEKPGGGTEITNYIVEKRESGTTAMOLVNS 22297
QY 248 ----- 247
Db 22298 VKRQIKVTHLTKYMEYSFRVSSNRFVSKPLESAPIIAEHPFVPPSAPTRPEVYHSA 22357
QY 248 ----- 247
Db 22358 NAMSIRWEEPYHDGGSKIIGYWEKKEKNTILWVKENKPCLECNKYKVTGLVEGLEYOFR 22417
QY 248 ----- 247
Db 22418 TYALNAGVSKASEASRPIAQNPDAPGREVTDVTRSTVSLIWSAPAYDGGSKVVGYYI 22477
QY 248 ----- 247
Db 22478 IERKPVSEVGDGRWLKCNYYTIVSDNFTVTALSEGDTYEFRLAKNAAGVISKSESTGP 22537
QY 248 ----- 247
Db 22538 VTCRDEYAPPKAELDARLHGLVITIRAGDLVLDAAVGGKPEPKIITWKGDKELDCEK 22597
QY 248 ----- 247
Db 22598 SLOVTKRRATAVIFCDRSDSGKYLTVKNASGTKAVSMVKVLDSPGCGKLTFSRVTO 22657
QY 248 ----- 247
Db 22658 EKCTLAWSLPQEDGAEITHYIVERRETSLRNWVIVEGECTLSYVVYTRLIKNNXYIFRV 22717
QY 248 ----- 247
Db 22718 RAVNKYGPVVESEPIVARNSETIPSPGIPBEVGTGKEHIIQWTKPESDGGNEISNY 22777
QY 248 ----- 247
Db 22778 LVDKREKKSRLWRVKNKYVYDTRLVKTVSLMEGCDYQFRVTAVNAAGNSEPSEASNFTIS 22837
QY 248 ----- 254
Db 22838 CREPSYTPGPPSAPRVVDTTKHSISLAWTKPMYDGGDTIVGYVLEMQEKDQDQWYRVHTN 22897
WLKAFYD
| | |

QY 255 ----- 254
Db 22898 ATIRNTEFTVPLKMGQKYSFRVAANVKGMEYSESIAETIEPVERIEIPDELADLKK 22957
QY 255 ----- 254
Db 22958 TVTIRAGASLRLMVSVSGRPPPVITWSKOGIDLASRAIIDTTESYSLLIVDKVNRDAGK 23017
QY 255 ----- 254
Db 23018 YTIEAENSGKKSATVLVKVYDTPGPCPSVKVKEVSRDSVTITWEIPTIDGAPVNNIV 23077
QY 255 ----- 254
Db 23078 EKREAAAMRAFKVTTKCKSKTLYRISGLVEGTWYFVRVLPENIYIGIGECETSDAFLVSEV 23137
QY 255 ----- 254
Db 23138 PLVPAKLEVVDTKSTVTLAWEKPLYDGGSLTGYVLEACKAGTERMKNVVTLKPTVLEH 23197
QY 255 ----- 254
Db 23198 TVTSLNBEQYLFRIIRAONEKGVSEPRETVTAVTVQDLRVLPTIDLSTMPOKTTHVPAGR 23257
QY 255 ----- 254
Db 23258 PVELVPIAGRPPPAASWFFAGSKLRESERVTVETHTKVAKLTIRETTIRDTGEVTLLEK 23317
QY 255 ----- 254
Db 23318 NVTGTTSETIKVIIIDKPGPPTGPIKIDEIDATISITISWEPPELGGAPLSGYVVEQRDA 23377
QY 255 ----- 254
Db 23378 HRCWLPVSESVTRSTFKFTRLTGEGNEYVFRVAATNRFSGSYLQSEVIECRSSIRIPGP 23437
QY 255 ----- 254
Db 23438 PETLIQIFDVSROGMTWLYPPEDDGGSOVTGYIVERKEVRADRWVRVKNKVPVTMYRST 23497
QY 255 ----- 254
Db 23498 GLTEGLEYHRVTAINARGSKPSRSPKPIVAMDPIAPPGKPNQPRVTDTRTSVSLAWS 23557
QY 255 ----- 254
Db 23558 VPEDEGSKVTGYLLEMQVDQHEWTKCNTPTPKIREYTLTHLPOGAERYFRVLACNAGG 23617
QY 255 ----- 260
Db 23618 PGEPAEVPGTVKVTMELEYDPYELDERYQEGIFVRQGGVIRLTIPIKGPFPICKWTKEG 23677
QY 261 ----- 260
Db 23678 QDISKRAMIATSETHFELVKEADRGDSGTVDLVLENKCGKAVYIKRVIGSPNSPEGP 23737
QY 261 ----- 260
Db 23738 LEYDDIQVRSVRVSWRPPADGGADILGYILERREVPKAAWYIDSVRGTSLVVGKLE 23797
QY 261 ----- 260
Db 23798 NVEYHFRVSAENQFISKPLKSEBPVTPKTLNPPEPPSPNPPEVLDVTKSSVLSWSRPK 23857
QY 261 ----- 260
Db 23858 DOGGSRTGYIYERKETSTDKVVRHNKTQITTTMYTGLVPOAEYQFRIIAQNDVGLSE 23917
QY 261 ----- 260
Db 23918 TSPASEPVVCKDPFDKPSQPGELEILISKDSVTLQWKEPCDGGKEILGYWVEYRQSGD 23977

Qy	245	-----	244
Db	19598	SGKTTITIAWKPFDGGAPITGYVEYKKSDDTDWKTSIOSLRCTEYITISGLITGAEVV	19657
Qy	245	-----	244
Db	19658	FRKSVNKVGASDPSDSDPOIAKEREERPLFDIDSEMRKTLIVKAGASFTMTVPFRGRP	19717
Qy	245	-----	244
Db	19718	VPNLWSKPDOTDLRTRAYVDTDSRTSLTIENANRNDGKYTLTIQNVLASAASLTVKV	19777
Qy	245	-----	244
Db	19778	LDTGPPNTITVQDVTKESAVLSWDPENDGGAPVKNYHIEKREASKANVSVTNNCRL	19837
Qy	245	-----	244
Db	19838	SYKVTNLQEGAIYFVRVSGENEFVGVI PAETKEGVKITEKPSPEKLGVTISIKSDVSLT	19897
Qy	245	-----	244
Db	19898	WLKPEHGGSRIVHYVVALEKGGOKMWKCAVAKSTHRVVSGLRENSEYFFRVFAENQAG	19957
Qy	245	-----	244
Db	19958	LSDPRELLPVLLIKEOLEPPEIDMKNFPSHTVYVRAGSNLKVDPISGKPLPKVTLSDRG	20017
Qy	245	-----	244
Db	20018	VPLKATRFNTEITAENLTINLKESVTADAGRYEITANSSTTKAFINIVVLDROPGPPT	20077
Qy	245	-----	244
Db	20078	GPVVISDITESVTLKWPPEKYDGGSQVTNYILLKRETSTAVTEVSATVARTMMKVMKL	20137
Qy	245	-----	244
Db	20138	TTGEYQFRIKAENRFGISDHDSACVTVKLPYTPPGPSTPWNTVNTRESITVQWHEPV	20197
Qy	245	-----	244
Db	20198	SNGSAVVGYHLEMKDRNSILWQANKLIVIRTHFKVTTISAGLIYFVRVYAENAAGVGK	20257
Qy	245	-----AFD-----	247
Db	20258	PSHPSEPVLAIDACEPPRNRVITDISKNSVLSWQQPAFDGGSKITGYIVERRDLPDGRW	20317
Qy	248	-----	247
Db	20318	TKASFTNVTEQFIISGLTQNSQYEFVRVARNVGSISNPSEVVGPICTIDSYGGPVIDL	20377
Qy	248	-----	247
Db	20378	PLEYEVVYKRGTSVKLRAGISGKPAPTIEWYKDDKELQTNALVCVENTTDLASILIKD	20437
Qy	248	-----	247
Db	20438	ADRLNSGCYELKLRNMGASATIRVQILDKPPGPGPIEFKVTVAEKITLLWRPPADDG	20497
Qy	248	-----	247
Db	20498	GAKITHYIVEKRETSRVVWSMVSEHLEBIIITTKIINKNEYIFRVRVKNYIGIGEPLES	20557
Qy	248	-----	247
Db	20558	DSVAKNAFVTPGPGPIPEVTKITKNSMTVVWSRPIADGSDISGYFLEKRDKSLGWFK	20617
Qy	248	-----	247
Db	20618	VLKETIRTRKVTGLTENSQYQYRVCVNAAGOGPSEPSEFYKAADPIDPPGPPAKIR	20677
Qy	248	-----	247

Db	20678	IADSTKSSITLGWSKPPVYDGSVAVTGYVVEIROGEEBEMTTVSTKGEVTRTEYVVVSNLKP	20737
Qy	248	-----	247
Db	20738	GVNTYFRVSAVNCAGQGEPIEMNEFPVQAKDILEAPEIDLVALRTSVIAKAGEDVQVLIP	20797
Qy	248	-----	247
Db	20798	FKGRPPPTVTRKDEKNLGSADARYSIENTDSSSLTTPQVTRNDTGKYLITENGVGEPK	20857
Qy	248	-----	247
Db	20858	SSTYSVKVLDTPAACQKLQVKHVSRGVTLLWDPPLIDGGSPIIINYVIEKRDATKRTWSV	20917
Qy	248	-----	247
Db	20918	VSHKCSSTFXLIDLSEKTPFFFRVLAENEIGIGPCETTEFPVKAEEVPAPIRDLMSKDS	20977
Qy	248	-----	247
Db	20978	TKTSVILSWTKPDDGGSVITEYVVERKKGGEQWTHAGISKTCEIEVSQLEQSVLEFR	21037
Qy	248	-----	247
Db	21038	VFAKNEKGLSDPVTIGPITVKELIITPEVDLSDIPGAQVTVRIGHNVHLELPYKGPKPS	21097
Qy	248	-----	247
Db	21098	ISWLKDGILPKESEFVRFSKTENKITLSIKNAKHEGSKYTVILDNAVCRVIAVITVITL	21157
Qy	248	-----	247
Db	21158	GPPSKPKGPIRFDEIKADSVILSWDVPDNGGGEITCYSIEKRETSQTNWKNWCVSVART	21217
Qy	248	-----	247
Db	21218	TFKVPNLVKAQYQFVRARENRYGVSQPLVSSIIIVAKHQFRIIPGPGPKVIYVNTSDGMS	21277
Qy	248	-----	247
Db	21278	LTWDAPVYDGGSEVTGFHVEKKERNISLWQKVNTSPISGREYRATGLVEGLDYQFRVAE	21337
Qy	248	-----	247
Db	21338	NSAGLSSPSPDKFTLAVSPVDPPTDYIDVTRETTILKNPPLRDGGSKIVGYSIEKR	21397
Qy	248	-----	247
Db	21398	QGNERNVRCNFTDVSECQYTVTGLSPGDRYEFRIIARNVAGTISPPSQSSGIIWTRDENV	21457
Qy	248	-----	247
Db	21458	PPIVEFGPEYFDGLIIKSGESLRIKALVQGRPVPRVTFKDGVEIEKRMMNEITDVLGST	21517
Qy	248	-----	247
Db	21518	SLFVRDATRDRHGVYTVVEAKNAGSAAEKVKVQDTPGKVGPPIRFTNITGKMTLWWD	21577
Qy	248	-----	247
Db	21578	APLNDGCAPITHYIIEKRETSRLAWALIEDKCEAQSVTAIKLINGEQFRVSAVNRFGV	21637
Qy	248	-----	247
Db	21638	GRPLDSDPVVAQIOYTVPDAPGIPSPNITGNSITLTWARPESDGGSEIOQYILERREKK	21697
Qy	248	-----	247
Db	21698	STRVVKVISKRPISRETKVTGLTEGNEYEFHVMAENNAGVGPASGISRLIKREPVNPP	21757
Qy	248	-----	247

Db	17378	DLRLVTLGLTEDHEYEFRVSAENAGVGEPSATVYKACDPVFKPGPPTNAHIVDTTKN	17437
QY	241	-----	240
Db	17438	SITLWANGKPIYDGGSEILGVVVEICKADEEWQIVTPQICLRYTRFEISKLTHEHQYKIR	17497
QY	241	-----	240
Db	17498	VCALNKVGLGEATSVPGTVKPEDKLEAPDLDELSELKGIIVRAGGSARIHIPFKGRPTP	17557
QY	241	-----	240
Db	17558	EITWSREGEFTDKVQIEKGVNTQLSIDNCDNRDAGKYLKLENSSGSKSAFVTVKVLVD	17617
QY	241	-----	240
Db	17618	TPGPPQNLAVKEVKDSAFVWEPPIIDGGAKVKNVIDKRESTRKAYANVSSKCSKTSE	17677
QY	241	-----	240
Db	17678	KVENLFEAIYPRVMAENEFVGVVPEVTVDAVKAABPPSPGKVTLTVDVSQTSASLMWE	17737
QY	241	-----	240
Db	17738	KPEHGGSRVLGVVEMQPKTEKWSIVAESKVCNAVVTGLSSGOEYQFVRKAYNEKGKS	17797
QY	241	-----	240
Db	17798	DPRVLGVPIAKDLTIQPSLKLFPNTYSIQAGEDLKEIPVIGRPRPNISWVKDGEPLAQ	17857
QY	241	-----	240
Db	17858	TTRVNVEETATSVLHIKEGNDKDFGKYVTATNSAGTATENLSVIVLEKPGPPVGVRRF	17917
QY	241	-----	240
Db	17918	DEVSADFVLSWEPPAYTGCQISNIVIVEKRDTTTTNHWMSATVARTTIKTKLTGTE	17977
QY	241	-----	240
Db	17978	YQPRIFAENRYKSAFLDSKAVIVQYPFKEPGPGPFVTSISKQMLVQWHEPVNDGCT	18037
QY	241	-----	240
Db	18038	KIIGYHLEQEKNSILWVKNLKTPIQDTRFKTTGLTDEGLEYEYFKVSAENIVGIGKPSKVS	18097
QY	241	-----	240
Db	18098	ECFVARDPCDPPORPEAIVITRNNVTWKWKPAYDGGSKITGYIVEKKDLDPGRWMAKF	18157
QY	241	-----	240
Db	18158	TNVLETEFTVSLGVEDQRYEVRVIARNAAGNFSEPSDSSGATARDEIDAPNASLDPKYK	18217
QY	241	-----	240
Db	18218	DVIIVHAGETFVLEADIRKPIPDVVVWSKDGELEETAARMEIKSTIQKTLVVKDCIRT	18277
QY	241	-----	240
Db	18278	DGGQYILKLSNVGGTKSIPITVKVLDROPPEGPKLVGTGTAECKCYLAWNPPLQDGGANI	18337
QY	241	-----	240
Db	18338	SHVIIKRETSRLSWTQVSTEVALNYKVTKLLPGNEYIFRVMVKNKYGIGEPLESGPYT	18397
QY	241	-----	240
Db	18398	ACNPYKPGPPSTPEYSAITKDSMVVTWARPVDDGGTEIEGYILEKRDKEGVRTKCNKK	18457
QY	241	-----	240
Db	18458	TLTDLRLVTLGLEHSHYEFVRVAENAGVGEFSEPSVIFYRACDALYPPGPPSNPKVTDT	18517
QY	241	-----	240
QY	241	-----	244
Db	18518	SRSSVSLAWSKPIYDGGAPVKGYVVEVKEAADEWTCPTPTGLQCKQTVIVKIKENTEX	18577
QY	245	-----	244
Db	18578	NFRICAINSEGVGEPATLPGSVVAQERIEBPPEIELDADLRKVVVLRASATLRLFVTIKGR	18637
QY	245	-----	244
Db	18638	PEPEVKWEKAEGILLTDRAQIEVTSSTMLVLDNVTRFDSGRYNLTLENNSSGSKTAFVNV	18697
QY	245	-----	244
Db	18698	VLDSPSPVNLITREVKKDSVTLWSWEPPLIDGGAKITNYIVEKRETRKAYAFITNCTK	18757
QY	245	-----	244
Db	18758	TFRIENLOEGCSYFPRVLASNEYIGLPAETTEPVKVSEPPPLPGRVTLVDVTRNTATI	18817
QY	245	-----	244
Db	18818	KWEKPESDGGSKITGYVVEWOTKSGSEKWSCTQVKTEATISGLTAGEEYVFRVAANEK	18877
QY	245	-----	244
Db	18878	GRSDPRQLGVPVIARDIEIKPSVELPFHTFNVKAREQLKIDVPFKGRQATVNRKDGOT	18937
QY	245	-----	244
Db	18938	LKETTRVNVSSKVTLSLSIKEAKEDVGTYELCVSNSAGSITVITIIVLDRPGPPGI	18997
QY	245	-----	244
Db	18998	RIDEVSCDSITISWNPPEYDGGCQISNYIVEKKETTSTTWHIVSQAVARTSIKIVRLTTC	19057
QY	245	-----	244
Db	19058	SEYQFRVCAENRYKGSYSSESSAVVAEYFPSPGPGTPKVVHATKSTMLVTWQVPVNDG	19117
QY	245	-----	244
Db	19118	GSRVIGYHLEYKERSILWSKANKILIAOTOMKVSGLDEGLMYERYVAENIAGIKGCSK	19177
QY	245	-----	244
Db	19178	SCEPVAPDPCDPPGQPEVTNITRKSLSKWSKPHYDGGAKITGYIVERRELDPGRWLKC	19237
QY	245	-----	244
Db	19238	NYTNIQETYFEVTELTEDQRYEFRVFARNAADSVSEPSSTGPIIVKDDVPEPRVMDVK	19297
QY	245	-----	244
Db	19298	FRDIVVVKAGEVLKINADIAGRPLVPISWAKDGIETEARTEIISTDNHTLLTVKDCIR	19357
QY	245	-----	244
Db	19358	RTGQYVLTKNVAGTRSVAVNCKVLDKPGPPAGPLEINGLTAEKCSLSWGRPOEDGGAD	19417
QY	245	-----	244
Db	19418	IDYIIVEKRETSHLAWTICEGELQMTSCKVTLLKKGNEYIFRVGTGVNKGVCGEPLESVAI	19477
QY	245	-----	244
Db	19478	KALDPETVPSPPTSLEITSWTKESMFLCWSRPESDGGSEISGYIIEERKNSLRWRVVK	19537
QY	245	-----	244
Db	19538	KPVYDLRVKSTGLREGCEYEYRVYAENAGLSLPSSETSLIRAEDPVFLPSPSPKPIVD	19597

QY 226 ----- 225
Db 15218 KVSGLTEGHEYEFRIMAENAGISAPSPSPFYKACDTVFKPGPGNPRVLDTSRSSISI 15277
QY 226 ----- 225
Db 15278 AWNKPIYDGGSEITGYMVEIALPEDEWQIVTPPAGLKATSYTTITGLTENQEKIRIYAM 15337
QY 226 ----- 225
Db 15338 NSEGLGEPALVPGTPKABDRMLPEIELDADLRKVVITIRACCTLRLEVPKIGRPAPEVKW 15397
QY 226 ----- 225
Db 15398 ARDHGESLDKASIBSTSSYTLIIVGNVNRFDGKYILTVENSCKSAFVNVRLDTPGP 15457
QY 226 ----- 225
Db 15458 PODLKVREVTKTSVTLTWDPPLLDGSKIKNYIVEKRESTRKAYSTVATNCHKTSMKWVQ 15517
QY 226 ----- PGK----- 228
Db 15518 LOEGCSYFRVLAENYEGIGLPAETAESVKASERPLPPGKITLMDVTRNSVLSWEKPEH 15577
QY 229 ----- 228
Db 15578 DGGSRILGYIVEMQTKGSDKWATCATVKVTEATITGLIQEEYSFRVSAONEKGISDPRQ 15637
QY 229 ----- 228
Db 15638 LSVPIAKDLVPPAPKLLFNTFTVLAGEDLKVDVPFGRPTPAVTWHKDNVPLKQTRV 15697
QY 229 ----- 228
Db 15698 NAESTENNSLLTIKDAREDVGHVYVVKLTNSAGEAIEFLANVILDKPGPPTGPKMDEV 15757
QY 229 ----- 228
Db 15758 ADSITLSWGPYPKYDGGSSINNYIVEKRDSTTTWQIVSAIVARTTIKACRLKTGCEYQFR 15817
QY 229 ----- 228
Db 15818 IAAENRYGKSTYLNSEPTVAQYFPKVPGPPTPVVTLSSRDSMEVOWNEPISDGGSRVIG 15877
QY 229 ----- 228
Db 15878 YHLERKERNILWVKLNKTPIPOTKFTTGLEGEVEYFRVSAENIVGIGKPSKVSECYV 15937
QY 229 ----- 228
Db 15938 ARDPCDPPRPEALIVTRNSVTLOWKKPTVDGGSKITGYIVEKKELPEGRWMAKSFNII 15997
QY 229 ----- 228
Db 15998 DTHFEVTGLVEDHRYEPRVIARNAAGVSEPSSESTGAIARDEVDPPRISMDPKYKDTIV 16057
QY 229 ----- 228
Db 16058 VHAGESFKVDADYGRPIPTIOWIKGQELSNARLEIKSTDPAFSLSVKDAVRVDSGNY 16117
QY 229 ----- 228
Db 16118 ILKAKNVAGERSVTNVKVLDRPGPPGPVVISGVTAECTLAWKPPLODGGSDIINYIV 16177
QY 229 ----- 228
Db 16178 ERRETSRLVTVVDANVQTLSCRVTKLLEGNEYTFRIMAVNKYGVGEPLSEPVAKNPF 16237
QY 229 ----- 228
Db 16238 VVPDAPKAPETVTKDSMTVWVERPASDGGSEILGVVLEKROKEGIRWTRCHKRLIGEL 16297
QY 229 ----- 228

Db 16298 RLRVTLGLIENHDYEFVSAENAGLSEPPSPSAYOKACDPIYKPPNPKNKVIDITRSSV 16357
QY 229 --DWLKAIFYD----- 236
Db 16358 FLSWSKPIYDGGCEIOGYIVEKCDVSYGWMTCTPPTGINKTNIEVEKLEKHEYNFRIC 16417
QY 237 ----- 236
Db 16418 AINKAGVGEHADVPPTIIVEEKLAPDIDLDELURKIINIRAGSRLFLFVPIKGRPTPEV 16477
QY 237 ----- 236
Db 16478 KWGKVDGEIRDAALIDVTSSTSLVLDNVNRYDSGKYTLTLENSGKTSAFVTVRVLDTDP 16537
QY 237 ----- 236
Db 16538 SPPVNLKVTETIKDSVSITWEPPLLDGGSKIKNYIVEKREATRKSAAVVTNCHKNWSKI 16597
QY 237 ----- KVAE----- 240
Db 16598 DQLOEGCSYFRVTAENYEGIGLPAQTADPIKVAEVPQPPGKITVDDVTRNSVLSWTKP 16657
QY 241 ----- 240
Db 16658 EHDGSKIIQYIVEMQAKHSEKARVKSLOAVITNLTOGBEYLFRVVAVNEKGRSDP 16717
QY 241 ----- 240
Db 16718 RSLAVPIAKDLVIEPDVKPAFSSYSVQGDCLKIEVPIISGRPKPTITWTKGDLPLKQTT 16777
QY 241 ----- 240
Db 16778 RINVTSDLTTLTUSIKETHKDDGGQYGITVANVVGOKTASIEIVTLDKPDPGPKVAFDD 16837
QY 241 ----- 240
Db 16838 VSAESITLSWNPPLYTGCGQITNYIVOKRDTTITVMDVVSATVARTLKVTCLKTGTGYQ 16897
QY 241 ----- 240
Db 16898 FRIFAENRYGQSFALSDPIVAQYYPYKEPGPGTPTFATAISKDSMVIOHHEPVNNGSPV 16957
QY 241 ----- 240
Db 16958 IGYHLERKERNILWTKVKNKTIITHDTQFKAQNLBEGIEYEFVYAENIVGVGKASKNSEC 17017
QY 241 ----- 240
Db 17018 YVARDPCDPPGTPEPIWVKRNEITLQWTKPVYDGGSMITGYIVEKRDLPDGRWMAKSFN 17077
QY 241 ----- 240
Db 17078 VIETQFTVSGLTQDQYEFVRVIAKNAAGAIKSPSDSTGPIITAKDEVELPRISMDPKFRDT 17137
QY 241 ----- 240
Db 17138 IVVNAGETFRLEADVHGKPLPTTIEWLRGDKIEESARCEIKNTDFKALLIVKDAIRIDGG 17197
QY 241 ----- 240
Db 17198 QYILRASNVAGSKSFPVNVKVLDRPGPEGPVQVGTSEKSLTWSPPLODGGSDISHY 17257
QY 241 ----- 240
Db 17258 VVEKRETSRLAWTVVASEVVVNTSLKVTKLLEGNEYVFRIMAVNKYGVGEPLSEAPVLMKN 17317
QY 241 ----- 240
Db 17318 PFVLPGPKPSLEVTNIAKDSMTVCWNRPDSDGGSEILGIYIVEKDRSGIRWIKCNKRIT 17377
QY 241 ----- 240

Db 12998 VISNITKDHMTVSWKPPADDDGGSPITGYLLEKRETOAVNMTKVNRKPIILRTLKATGLQE 13057
QY 209 ----- 208
Db 13058 GTEYEFRTAINKAGPKPSDASKAAYARDPOYPAPAPAFKPYDTTRSSVLSWKGKPAY 13117
QY 209 ----- 208
Db 13118 DGGSPILGYLVEVKRADSDNWRCLNQLONLQKTRFETGLMEDTOYQFRVAVNKGISD 13177
QY 209 ----- 208
Db 13178 PSDVPDKHPKDLIPPEGELDADLRKLILRAGVTMRLYVPVKGRPPKTIWSPKNVNL 13237
QY 209 ----- 208
Db 13238 RDRIGLDIKSTDFRFLRCENVNKYDAGKYILTLNCSGKKEYTIWVKVLDTPGPPVNT 13297
QY 209 ----- 208
Db 13298 VKEISKDSAYVTWEPPIIDGGSPIIYVQKRAERKSWSTVTTECSKTSFRVANLECK 13357
QY 209 ----- 208
Db 13358 SYFFRFAENEXYIGDPTGTRDAVKASQTPGPVVDLKVRSVSKSSCSIGWKKPHSDGGSR 13417
QY 209 ----- 208
Db 13418 IIGYVDFUTEENKQVRMKSLLOYSAKDLTEGREYTFRVAENENGEGTPTSEITVVAR 13477
QY 209 ----- 208
Db 13478 DDVAPDLDLGLDPLCYLAKENSFRKIPIKGPAPSVSWKGEDPLATDTRVSVSS 13537
QY 209 ----- 208
Db 13538 AVNTTLIVYDCQSDAGKYITLKNVAGTKEGTISIKVVGKPGIPTGPIKDFDEVTAEMT 13597
QY 209 ----- 214
Db 13598 LKWAPPKDDGGSEITNYILEKRDVNNKWTCASAVOKTTFRVLRLHEGMEYTFRVAEN 13657
QY 215 ----- 214
Db 13658 KYGVEGLKSEPIVARHPFDVPDAPPPNIVDVHDSVSLTWTDPKKTGGSPITGYHLEF 13717
QY 215 ----- 214
Db 13718 KERNSLLMKRANKTPIRMDFKVTGLTEGLEVEFRVMAINLAGVGKPSLPSEPVVALLDPI 13777
QY 215 ----- 214
Db 13778 DPPGKPEVINITRNSVTLIWTBPKYDGGHKLTYIVEKRDLPKSWMKANHVNPECAFT 13837
QY 215 ----- 214
Db 13838 VTDLVEGGYEFPIRAKNTAGISAPESTETIICKDEYEAPTIWLDPTIKDGLTIKAGD 13897
QY 215 ----- 214
Db 13898 TIVLNAISTLGLKPLKSSWSKAGDIRPSDITQITSTPTSSMLTIKYATRKDAGEYTTVA 13957
QY 215 ----- 214
Db 13958 TNPFGTKVEHKVTVLDVPGPGPVEISNVAEKATLTWTPPLEDGGSPIKSYILEKRET 14017
QY 215 ----- 217
Db 14018 SRLMTVVSEDIQSCRHVATKLIQNEXYIFRVSANVHYKGPVQSEPKVMDRFGPPG 14077
QY 218 ----- 217
Db 14078 PEKPEVSNVTKNTATVSNKRPVDDGSGSEITGYHVERREKKSRLRWRAIKTPVSDLRCKVT 14137

QY 218 ----- TOKSLSLS-- 225
Db 14138 GLQEGSTYEFVRYSAENRAGIGPPSEASDVLKMDAAYPGPPSPNHVTDTTKKSASLAWG 14197
QY 226 ----- 225
Db 14198 KPHYDGLLEITGYVVEHQKVGDSEAWIKDITGTALRITQFVVVDLQTKERYNFRISAINDA 14257
QY 226 ----- 225
Db 14258 GVGEPAVIDDVEIVEREMADPDELRLTLVVRAGLSIRIFVPPIKGRPAPEVTTWKDN 14317
QY 226 ----- 225
Db 14318 INLKNRANIENTESFTLLIIPECNRYDTGKFVMTIENPAGKSGFVNVRLDTPGPVLNL 14377
QY 226 ----- 225
Db 14378 RPTDITKDSVTLHWDLPLIDGGSRTNYIVEKREATRKSYSTATTKCHKTYKVTGLSEG 14437
QY 226 ----- 225
Db 14438 CEYFFRYMAENEXYIGCEPTEPTEPVKASEAPSPDPSLNIMDITKSTVSLAWPKPHDGG 14497
QY 226 ----- 225
Db 14498 KITGVVIEAQRKGSQDOWHITTVKGLCEVVRNLTEGEYTFQYMAVNSAGRSAPRESRPV 14557
QY 226 ----- 225
Db 14558 IVKEQTMPLPELDLRGIYQKLVIAKAGDNIKVEIPVLRPKPTVTWKKGQOILKQTORVNF 14617
QY 226 ----- 225
Db 14618 ETTATSTILNINECVRSDSGPYPLTARNIVGEVDVITIQVHDI PGPTGPIKDFEVSSD 14677
QY 226 ----- 225
Db 14678 FVTFSDPPDPPDGGVPISNYVVMRQDSTTWVELATTVIRTYKATRLTTGLQYQFRVK 14737
QY 226 ----- 225
Db 14738 AQNRXYGVGPGITSACIVANYPFKVPPOPTQVTAVKDSMTLSWHEPLSDGSGSPILGVH 14797
QY 226 ----- 225
Db 14798 VERKERNGILWQTVSKALVPGNIFKSSGLTDGIAVEFRVIAENMACKSKPSKPEPMAL 14857
QY 226 ----- 225
Db 14858 DPIDPPGKPVPLNITRHTVTLKWAPEYTGFGKITSYIVEKRDLPNGRMLKANFSNILEN 14917
QY 226 ----- 225
Db 14918 EFTVSGLTEDAAAYEFVRIAANAAGALSPPSEPSDAITCRDDVEAPKIKVDVKFKOTVILK 14977
QY 226 ----- 225
Db 14978 AGEAFRLEADVSGRPPTMWSKDGKEGTAKLEIKIADFTNLVKNKDSRDSGAYTL 15037
QY 226 ----- 225
Db 15038 TATNPGGFAKHIFNVKVLDRPGPPEGLAVTEVTSEKCVLSWFPPLDDGGAKIDHYIVOK 15097
QY 226 ----- 225
Db 15098 RETSLANTVNASEVOVTKLVTKLLKGNEYIFRVMVANKYGVGEPISEPLEVLANPYGP 15157
QY 226 ----- 225
Db 15158 PDPPKNPEVTTITKDSMVVWCHGHPDSDGSGSEIINIVERRDKAGORWIKCNKKTLDLRY 15217

```
QY 177 ----- 176
Db 10838 DGGSVITNVYERRDVASQWSPSATSKKSHFAKHLNEGQYLFRVAENQYGRGPF 10897
QY 177 ----- 176
Db 10898 VETPKPIKALDPLHPGPKDLHHVDVDKTEVSLWNKPKDRDGGSPITGYLVEYQEBGTQ 10957
QY 177 ----- 176
Db 10958 DWIKFKTVNLECVVTGLQOGKTYFRVKAENIVGLGLPDTTPIECOEKLVPPSVELDV 11017
QY 177 ----- 176
Db 11018 KLIEGLVVKAGTVRFPAILIRGVPVPTAKWTTDGESEIKTDEHVTVEFDNFSSVLITKNCL 11077
QY 177 ----- 176
Db 11078 RRDTEYOITVSNAGSKTVAVHLTVLDVPGPPTGINILDVTPEHMTISWQPPKDDGGS 11137
QY 177 ----- 176
Db 11138 PVINIVEKODTRKDTGWVSSGSSKTKLIPHLOKCEYVFRVRAENKICVGPPLDSTP 11197
QY 177 ----- 176
Db 11198 TVAKHKFPSPPGKPVVTDITENAAVTSWTLPKSDGSGSPITGYMERREVTKWVRVVK 11257
QY 177 ----- 176
Db 11258 TPIADLKFRVTGLYEGNTYFRVFAENLAGLSKPSDDPIKACRPKPPGPPINPKLKD 11317
QY 177 ----- 176
Db 11318 KSRETADLVWTKPLSDGSGSPILGYVVEVCQPGTAQWNRINKDELIROCAFRVPGLEGNE 11377
QY 177 ----- 176
Db 11378 YRFRIKAANIVGEGEPRELAESVIATKDLHPPEVELDVTCRDVITVRVGOTIRILARVK 11437
QY 177 ----- 176
Db 11438 RPEPDITWTKRGLVREKRVLDLIQDLPRVELQIKEAVRADHGKYYIISAKNSSGHAQGSA 11497
QY 177 ----- 176
Db 11498 IVNVLDPRGPCQNLKVTNVTKENCTISWENPLDNGGSEITNFIVEYRKNQKGSIVASD 11557
QY 177 ----- 176
Db 11558 VTKRLIKANLANNEYFRVCAENKVGVGPTIETKTPILAINPIDRGPENLHIADKGK 11617
QY 177 ----- 176
Db 11618 TFVYLKWRDPDYDGGSPNLSYHVERRLKGSDDWERVHKSGIKETHYMWDRVCVENQIYEFR 11677
QY 177 ----- 180
Db 11678 VQTKNEGGSWVKTEEVVVKEDLQKPVLDLKLGSGLTVKAGDTIRLEAGVRGKPFPEVA 11737
QY 181 ----- 193
Db 11738 WTKDKDATDLTRSPRKIDTRADSSKFSLTAKRSDGGKYVVTATNTAGSFVAYATVNVL 11797
QY 194 ----- 193
Db 11798 DKPCPVRLKIVDVSSDRCTVCWDPDDEGGEIQNYILEKETKRWVWSTYSATVLTGP 11857
QY 194 ----- 195
Db 11858 TTVTRLIEGNEYIFRVAENKIGTPPTESKPVIAKTKYDKRGRDPPPEVTKVSKKEMTV 11917
QY 196 ----- 195
```

```
Db 11918 VWNPEYDGGKSGITGYFLEKKEKHSTRWVPVNKSAIPERRMKVONLLPDHEYOFRVKAEN 11977
QY 196 ----- 195
Db 11978 EIGIGEPSLPSRPVAKDPIEPPGPTNFRVDTTKHSITLGMGKPYVDGAPIIGYVVE 12037
QY 196 ----- 195
Db 12038 MRPKIADASPEGKRCNAAAQLVRKEFTVTSLDNEQEYFRVCAQNOVGIGRPAELKEA 12097
QY 196 ----- 195
Db 12098 IKPEILEPPEIDLDASMRKLIVIRAGCPIRLFAIVRGRPAKVTWRKVGIDNVVRKGQV 12157
QY 196 ----- 195
Db 12158 DLVDTMAFLVIPNSTRDDSGKYSLLTVNPAGKAVFVNVVRVLDTPGPVSDLKVSVDVTKTS 12217
QY 196 ----- 195
Db 12218 CHVSWAPPENDGGSQVTHYIVKREADRKTWSTVTPVKKTSFHVNTNLVPGNEYFRVTA 12277
QY 196 ----- 195
Db 12278 VNEYGCVPTDVPKPVLASDPLSEPDPPRKLEVTMTKNSATLAWLPLRDGGAKIDGYI 12337
QY 196 ----- 195
Db 12338 TSYREEEQPADRWTEYSVWKDLSLVVTGLKEGKKYFRVAARNAVGSLPREAEGVYEAK 12397
QY 196 ----- 202
Db 12398 EQLLPPKILMPEOITTKAGKKLRIEBAHVYCKPHPTCKKKGGEDEVTSSSHAVHKADSSS 12457
QY 203 ----- 202
Db 12458 ILIIKDVTBKDSGYSLTAENSSGTDQIKIVVMDAPGPPQPFDISDIDADACSLSWH 12517
QY 203 ----- 202
Db 12518 IPLEDGGSNITNYIVEKCDVSRGDWVTALASVTKTSCRVGKLIPGQYIFRVAENRFGI 12577
QY 203 ----- 202
Db 12578 SEPLTSPKVAQPFPGVSEPKNARVTKVNKDCIFVANDRPDSDGSGSPIIGYLIERKERN 12637
QY 203 ----- 202
Db 12638 SLLWVKANDFLVRSTEYPCAGLVEGLEYSFRIYALNKGSSPPSKPTEYVTARMPVDPG 12697
QY 203 ----- 202
Db 12698 KPEVIDVTKSTVSLIWARPKHGGSKIIGYFVEACKLPQDKWVRNCTAPHQIPQEEYAT 12757
QY 203 ----- 202
Db 12758 GLEEKAOYOFRAIARTAVNISPPSEPSDPVTILAENVPPIRLSVAMKSLTLTKAGTNVC 12817
QY 203 ----- 202
Db 12818 LDATVFGKPMPTVSWKKDGTLLKPAEGIKWAMQNLCTLELFSVNRKDSGDTITAE NSS 12877
QY 203 ----- 202
Db 12878 GSKSATIKLVLDKPGPPASVKINKMYSDRAMLSWEPPELDEGGEITNYIVDKRETSRPN 12937
QY 203 ----- 208
Db 12938 WAQVSATVPITSCVSEKLEHGYQFRICAENKYGVGDPVFTPEATAKNPDYDPRGRCDPP 12997
QY 209 ----- 208
```

```

RESULT 15
Q8WZB3
ID Q8WZB3 PRELIMINARY; PRT; 26926 AA.
AC Q8WZB3;
DT 01-MAR-2002 (TEMBLrel. 20, Created)
DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE N2B-titin isoform.
GN TTN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=20309627; PubMed=10850961;
RA Freiburg A., Trombittas K., Hell W., Cazorla O., Fougereousse F.,
RA Centner T., Kolmerer B., Witt C., Beckmann J.S., Gregorio C.C.,
RA Granzier H., Labeit S.;
RT "Series of exon-skipping events in the elastic spring region of titin
RL as the structural basis for myofibrillar elastic diversity.";
RL Circ. Res. 86:1114-1121(2000).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=21573839; PubMed=11717165;
RA Bang M.L., Centner T., Fornoff F., Geach A.J., Gotthardt M.,
RA McNabb M., Witt C.C., Labeit D., Gregorio C.C., Granzier H.,
RA Labeit S.;
RT "The complete gene sequence of titin, expression of an unusual ~700
RT kDa titin isoform and its interaction with obscurin identify a novel
RT z-line to I-band linking system.";
RL Circ. Res. 89:1065-1072(2001).
DR EMBL: AJ277892; CAD12455.1; -
DR InterPro: IPR000282; Cytok_receptor_2.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR000577; FGGY_kin.
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR001092; HLH_basic.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003598; Ig_c2.
DR InterPro: IPR003600; Ig_like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR002016; Peroxidase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00041; fn3; 132.
DR Pfam: PF00047; Ig; 91.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00060; FN3; 132.
DR SMART: SM00409; IG; 111.
DR SMART: SM00408; IGC2; 93.
DR SMART: SM00410; IG_like; 3.
DR SMART: SM00220; S_TKc; 1.
DR SMART: SM00219; TyKc; 1.
DR PROSITE: PS00933; FGGY_KINASES_1; UNKNOWN_1.
DR PROSITE: PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
DR PROSITE: PS00435; PEROXIDASE_1; UNKNOWN_1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
SQ SEQUENCE 26926 AA; 2993285 MW; 169AB42637A7C1FB CRC64;

Query Match 34.9%; Score 466; DB 4; Length 26926;
Best Local Similarity 0.6%; Pred. No. 0.00029;
Matches 155; Conservative 48; Mismatches 40; Indels 26150; Gaps 46;

QY 3 KTHTCPP----- 9
| | |
| | |
Db 98 KAETAPNFVORLQSMVROGSOVRLOVRVTGIPTPVVKFYRDGAETQSLDFQISOEGD 157
QY 10 ----- 9

```

Db 1238 SESEAVESGFDLRKIKNYRILEGMYTFHCKMSGYPLPKIAWYKDGKRRIKHGERYQMDFLQ 1297
QY 34 ----- 33
Db 1298 DGRASLRIPVLPEDRGYITAFASNIKGNACISGKLYVEPAAPLGAPTYIPTLEPVSRIR 1357
QY 34 -----ISRTP----- 38
: || :
Db 1358 SLSPRSRSRPTIRMSPARMSPARMSPARMSPARMSGRLEETDESOLERLYKPVFLKP 1417
QY 39 ----- 38
Db 1418 VSFKLEGOTARFDLKVGRPMPTFWFHDGQOIVNDYTHKVIKEDGTSQSLIIVPATPS 1477
QY 39 ----- 38
Db 1478 DSGEWTVAQNRAGRSSISVILTVEAVEHOVKPMFVEKLNKYNIKEGSQLEMKVRATGNP 1537
QY 39 ----- 38
Db 1538 NPDIVLKNSDIIVPHKPKIRIEGTGKEALKIDSTVSQDSAWYTATAINKAGRDTTRC 1597
QY 39 ----- 38
Db 1598 KVNVEVEFAPEPERKLIIPRGTYRAKETAAPELEPLHLRYGOEOWEGDLYDKEKQKRP 1657
QY 39 ----- 38
Db 1658 FFKKKLITSLRKRFPAHECRLTPIGPTMVVWHLHDGKPLEAANRLMINEFGYCSLD 1717
QY 39 -----EUTC----- 42
: ||
Db 1718 YGVAYSRDGIIITCRATNKYGTHTSATLIVKDEKSLVESQLPEGRKGLRIEELERMA 1777
QY 43 ----- 42
Db 1778 HEGALTGVTTDOKEKQKPDIVLYPEVRVLEGETARFRVRTGYPOPKNWYLNGLIRK 1837
QY 43 ----- 42
Db 1838 SKRFRVRYDGIHYLDIVDCSYDTGEVKVTAENPEGVIEHKVLEIQOEDFRSVLRAP 1897
QY 43 ----- 42
Db 1898 EPRPEFHVHPGKLOFEVQKQVDRPVDTTETKEWKLKRAERITHEKVPESBELSKFKR 1957
QY 43 ----- 42
Db 1958 RTEEGYEAITAVELKSRKKDSYBELLRKTKDELLHMTKELTEBEKKALAEKGKITIPT 2017
QY 43 ----- 42
Db 2018 FKPDKIELSPSEAPKIFERIOSQTVGQSDAHFRVRVVGKDPDPECWYKNGVKIERSDR 2077
QY 43 -----VVVDVSHED----- 51
: || :
Db 2078 IYWYWPEDNVCVIRDTVAEDSASIMVKAINIAGETSSHAPLLVOAKLITFTQELQDV 2137
QY 52 -----PEVKFNWYDGEVH----- 66
: || :
Db 2138 VAKEKDTMATFECETSEPPFVKYKWKDGMVEHGDYKRMHSDRKVHFLSILTIDTSDAED 2197
QY 67 -----NAKT----- 70
: ||
Db 2198 YSCLVEDENVKTTAKLIVAGAVEFVKELQDIEVPSYSGELECIVSPENIEGKWYHND 2257
QY 71 -----KPR----- 73
: || :
Db 2258 VELKSNKYTITSRGRONLTVKDVTKEDQGEYSFVIDGKKTCKLKMKPRPIALQGLS 2317
QY 74 ----- 73
Db 2318 DQKVCEDIVLEVKVSLESVEGVWMDQGEVQPSDRVHIIVIDKOSHMLLIEDMTKEDAG 2377

QY 74 ----- 73
Db 2378 NYSFTIPALGSLTSGRVSVYSDVITPLKDVNVIETRAVLECKVSPDVTSVKWLND 2437
QY 74 ----- 73
Db 2438 QIKPDDRVOAIVKGTQORLVINRTHASDEGPYKLIIVGRVETNCNLSVEKIKIIRGLDIT 2497
QY 74 ----- 73
Db 2498 CPTONVVFEVLSHSGIDVLNFKDKEIKPSSKYKIEAHGKIYKLTVLNMMKDDGKVT 2557
QY 74 ----- 73
Db 2558 FYAGENITSGKLTVAGGAIKPLTDQTVAESQEAPECEVANPDSKGEWLRDGHPLTN 2617
QY 74 ----- 73
Db 2618 NIRSESDGHKRLIIAATKLDIGEYTKVATSKTSAKLKVEAVKIKKTLKNLTVTETQD 2677
QY 74 ----- 73
Db 2678 AVFTVELTHPNVGVQWIKNGVLESNEKYAISVGTIYSLRIKNCIAIVDESUYGFLGR 2737
QY 74 ----- 73
Db 2738 LGASARLHVEITVKIIPKPDVTALENATVAFESVSHDTPVVKWFHKNVEIKPSDKHLV 2797
QY 74 ----- 73
Db 2798 SERKVHKLMLQNISPDAGEYTAVVQLECKAKFVETLHITKMKNIEVPETKTASPEC 2857
QY 74 ----- 73
Db 2858 EYSHFNVPMSMLKNGVEIEMSEKFIIVOGKLUHOLIIMNTSTEDSAEYTFVCGNDQVSAT 2917
QY 74 ----- 73
Db 2918 LTVTIMITSMLKDINAEEKDITTEFVTVNYESISYKWLKNGVEIKSTDCKOMRTKKLTH 2977
QY 74 ----- 73
Db 2978 SLNIRNVHFGDAADYTFVAGKATSTATLYVEARHIEFRKHIDKIVLEKKRAMPECEVSE 3037
QY 74 ----- 73
Db 3038 PDITVQWKKDDQELQITDRIKIQEKYVHRLIIPSTRMSDAGKYTVVAGGNVSTAKLFVE 3097
QY 74 ----- 73
Db 3098 GRDVRIRSIKKEVQVIEKQRAVVEFEVNEDDVDHMYKDGIEINFQVQERHKYVVERRIH 3157
QY 74 ----- 73
Db 3158 RMFISETRQSDAGEYTFVAGNRSSVTLYVNAPEPPQVLQELQPVTVQSGKPARFCAVIS 3217
QY 74 -----EE----- 75
: ||
Db 3218 GRPOPKISWYKEEQLLSTGFKCKFLHDGOEYTLILLIEAPPEDAAYTCEAKNDYGVATPS 3277
QY 76 ----- 75
Db 3278 ASLSVEPEVVSPOEMPVYPPAIITPLQDITVTSQOPARFQCRVSGTDLKVSWSKDKK 3337
QY 76 -----QYNSTYR----- 83
: || :
Db 3338 IKPSRFRMTQFEDTYQLEIAEAYPEDEGTFTFVASNAVGQVSSSTANLSLEAPESILHER 3397
QY 84 ----- 83
Db 3398 IEQEIEMEMKEFSSFLSAEBEGLHSAELQLSKINETLELLESPPVYSTKFDSEKEGYGP 3457

QY 84 -----VSVL----- 87
Db 3458 IFIKEVSNADISMGDVATLSVTVIGIPKPKIQWFFNGVLLTPSADYKRFVFDGDDHSLIIL 3517
QY 88 ----- 87
Db 3518 FTKLEDEGEYTCMASNDYKGTICSAYLKINSKGEGCHKDTETESAVAKSLEKLGPCPPHF 3577
QY 88 -----TVLH----- 91
Db 3578 LKELKPIRCAOGLPAIFEYTVWGPAPTIVTFWKENKOLCTSVYVYTHHNPNGSGTFVIND 3637
QY 92 ----- 91
Db 3638 POREDSGLYICKAENMLGESTCAAELLVLEDTDMTPCKAKSTPEAPEDFPOTPLKGP 3697
QY 92 ----- 91
Db 3698 AVEALDSEGEIATVKTILKAALITEENQOLSIEHIAKANELSSQLPLGAELQSLIEQ 3757
QY 92 ----- 91
Db 3758 DKLTPESTREFLCINGSIHQPLKEPSPNLQLOIIVQSQKTFSEKGIILMPPEPETOAVLSD 3817
QY 92 ----- 91
Db 3818 TEKIFPSAMSTEQINSLTVEPLKTLAEPEGNYPOSSIEPPMHSYLTSAEEVLSPEKKT 3877
QY 92 ----- 91
Db 3878 VSDTNREQRVTLOKQEAQSALILQSIAEGHVESLOSQPDVNIQVNYEPLVPSEHSCTEG 3937
QY 92 ----- 91
Db 3938 GKILLIESANPLENAGDSAVRIEBSKLSRPLALEEKQVLLKEHSDNVVWPPDQIIIESK 3997
QY 92 -----QDWL- 95
Db 3998 REPVAIKKVOEVOGRDILLSKESLISGIPERQRLNLKIQICRALQAQAAVASEQPLFSEWL 4057
QY 96 ----- 95
Db 4058 NIEKVEAVNITQEPHIMCMYLVTSAKSVTEEVTHIIEDVDPQMANLKMELRDALCAI 4117
QY 96 ----- 95
Db 4118 IYEEIDILTAEGPRIQOQAKTSLQEENDSFGSKQKVEPITEPEVESKYLISPEVSYFNV 4177
QY 96 ----- 95
Db 4178 QSRVKYLDATPVTKGASAVVSEKODESLKPSKEESESSESCTEEVATVKIQEABGGF 4237
QY 96 ----- 95
Db 4238 IKEDGPMIHTPLVDTVSEGDIVHLTTSITNAKEVNWYFENKLVPSDEKPKCLODQNTYT 4297
QY 96 -----NGK----- 98
Db 4298 LVIDKVNTEDHQGYVCEALNDSKTTATSAKLTVVKKRAAPVKKRIEPLVALGHLAKFT 4357
QY 99 -----EYKCKVSN----- 106
Db 4358 CEIOSAPNVRQWFKAGRIEYSDKCSIRSISKYISSLEILLRTQVVDGGEYTCRASNEYGS 4417
QY 107 ----- 106
Db 4418 VSCATLTVTVPGEKKVRKLLPERKPEKVEVVLKSVLKRPEEPEKPEKLEKVKK 4477
QY 107 ----- 106
Db 4478 PAVPEPPPPKVEVEVPTVTKRKRIPEPTKVPEIKPAIPLPAPEPKPKPEAEVKTIPK 4537
QY 107 -----KALPAPIE----- 114

Db 4538 PPVEPEPTPIAAPVTVPVGGKAEAKAPKEAAKPGPIKGVPKKTSPIAEERRKLPG 4597
QY 115 ----- 114
Db 4598 SGGEKPPDEAFYTYQLKAVPLKFVKEIKDIIILTESEFVGSSAIFECVLSPSTAITTWMKD 4657
QY 115 ----- 114
Db 4658 GSNIRESPKHRFIADGDKDKLHIIIDVQLSDAGEYTCVLRGNKKEKTSTAKLVVEELPVRF 4717
QY 115 -KTIS----- 118
Db 4718 VKTLEEVTVVKGQPLYLSCNELKRDVWRKDGKIVVEKPGRIVPGVIGLMRALJTINDA 4777
QY 119 ----- 118
Db 4778 DDTDAGTYTVTVENANNLECSVCVVVEVIRDMVLVKPIRDQHVKPKGTAFACDIADKTP 4837
QY 119 ----- 118
Db 4838 NIKWFKGYDEIPAPBNDKTEILRDGNHLYLKIKNAMPEDIAEYAVEIEGKRYPAKLTLGE 4897
QY 119 ----- 118
Db 4898 REVELLKPIEDVTIYKESASFDAEISEADIPGQWKLKELLRPSPTCEIKAEGGKRFLT 4957
QY 119 ----- 118
Db 4958 LRKVKLDQAGEVLYQALNAITTAITLVKEIHELDFAVPLKDVTVPERQARFECVLTREAN 5017
QY 119 ----- 118
Db 5018 VIKSGPDIKSSDKPDIADGKKHILVINDSQFDDGCVYTAEVGKTSARLFTVGTGRL 5077
QY 119 ----- 118
Db 5078 KFMSPLEDQTVKEGETATFVCELSHEKMHVWFKNDAKLHTSRTVLISSEKTHKLEMK 5137
QY 119 ----- 118
Db 5138 VTLLDISQIKAQVKELSSTAQLKVLEADPYFTVKLHDKTAVEKDEITLKCEVSKDVPVKW 5197
QY 119 ----- 118
Db 5198 FKDGEEIVSPKYSIKADGLRRLIKKADLKDKEVYVCDGCTDKTKANVTVEARLIKVE 5257
QY 119 -----KAGQP----- 124
Db 5258 KPLYGVEFVGETAHFEIELSEPDVHGQWKLKGPLTASPDCEIIEDGKKHILILHNCQL 5317
QY 125 -----REPO----- 128
Db 5318 GMTGEVSFOANAKSAANLKVKEPLIFITPLSDVKVFEXDEAKFECEVSREPRTFRWLK 5377
QY 129 ----- 128
Db 5378 GTQEITGDDRFELIKDGTKHSWIKSAAFEDKAYMEAEKHTSGKLIIEGIRLKFITP 5437
QY 129 ----- 128
Db 5438 LKDVTAKEKESAVFTVELSHDNIRVKWFKNDQRLHTTRSVSMQDEGKTHSITFKDLSIDD 5497
QY 129 ----- 128
Db 5498 TSQIRVEAMGMSSEAKLTVLEGDPYFTGKLQDYTGVEKDEVILQCEISKADAPVKWPKDG 5557
QY 129 ----- 128
Db 5558 KEIKPSKNAVIKADGKKRMLILKALSDIGQYTCDCGTDKTSGLKDIEDREIKLVRPLH 5617
QY 129 ----- 128

Db 5618 SVEVETETARFETEISEDDIHANWKLGEALLQTPDCEIKERKTHLSVLHNCRLDQTG 5677
QY 129 ----- 128
Db 5678 GVDFOANVKSALHVRKPRVIGLLRPLKDVTVTAGETATFDCELSYEDIPVEWYLGKK 5737
QY 129 ----- 128
Db 5738 LEPSDKVVRSEKGVHTLRLDVKLEDAGEVQLTAKDFKTHANLFVKEPPVEFTKPLEDQ 5797
QY 129 ----- 128
Db 5798 TVEGATAVLECEVSRENAKVFKNGTEILKSKKYEIVADGRVKLVIHDTCPEDIKY 5857
QY 129 ----- 128
Db 5858 TCDAKDFKTSCLNVPPHVEFLRPLTDLOVREKEMARFECELSRENKVKFKDGAETK 5917
QY 129 ----- 128
Db 5918 KGKKYDIISKGAVRIILVINKCLLDDAEYSCEVRTARTSGMLTVLEEEAVFTKNLANIEV 5977
QY 129 ----- 128
Db 5978 SETDTIKLCEVSKPGAIEVIWYKGDEEIIETGRYEILTEGRKRIILVIONAHLEDAGNYNC 6037
QY 129 ----- 128
Db 6038 RLPSSRTDGKVKVHELAAEFISKQNLLEGEKAEFVCSISKESPPVQWKRDDKTLESG 6097
QY 129 ----- 128
Db 6098 DKYDVIADGKRVLVVKDATLQDMGTGVVYVVGAAARAAHLTVIEKLRIVVPLKDTFVKQ 6157
QY 129 ----- -VYTL- 132
Db 6158 QEYVFNCEVNTGAKAKWPNBEAIPDSSKYIILQKDLVYTLRIADHLDQOANYSLT 6217
QY 133 ----- |||| 132
Db 6218 NHRGENVKSANLIVEEDLRIVEPLKDIETWEKKSVTFWCKVNLNVLTKWTKNGEVP 6277
QY 133 ----- 132
Db 6278 FDNRSYRVDKYKHLMTIKDCGFPDGEYIVTAGQDKSVAELLIIEAPTEFVEHLEDQV 6337
QY 133 ----- 132
Db 6338 TEFDDAVFCQLSREKANKWYRNGREIKEGKKYKFKDGSIHRLTIKDCRLDDECEYAC 6397
QY 133 ----- 132
Db 6398 GVEDRSARLRFVEETPVEIIRPPQDILEAPGADVVFLELNKDKVEQWLNNMVVVG 6457
QY 133 ----- 132
Db 6458 DKHQMSECKIHLQICDIKPRDGEYRFIANDKEARAKLELAAPKIKTADQDLVDVG 6517
QY 133 ----- 132
Db 6518 KPLTWVVPYDAYPKAAEWFKENEPLSTKTIDTTABOTSFRILEAKKGDKGRYKIVLQNK 6577
QY 133 ----- 132
Db 6578 HGKABGFINLVIDVPGPVNLEVTFDGEVSLAWEEPLTDGGSKIIGYVVERDIRK 6637
QY 133 ----- 132
Db 6638 TWVLTADRAESCEFTVTGLQGGVEYLFVRVARNRVGTGEPVETDNPVEARSKYDVPGP 6697
QY 133 ----- 132
Db 6698 LNVITDVNRFGVSLTWPEPPYDGGAEITNVVIELRDKTSIRWDTAMTVRAEDLSATVTD 6757

QY 133 ----- 132
Db 6758 VVEGOEYSPRVAQNRIQVCKPSAATPFVKVADPIERPSPPNLTSSDOTQSSVQLKWP 6817
QY 133 ----- 132
Db 6818 PLKGGSPILGYIIERCCEGKDNWIRCNMKNLVPULTYKVTGLEKGNKYLYRVAENKAGV 6877
QY 133 ----- 132
Db 6878 SDPSEILGLTADDAFVEPTMDLSAFKDGLEVIVPNPITILVPSTGYPRPTATWCFGDKV 6937
QY 133 ----- 132
Db 6938 LETGDRVKMKTLSAYAEVLISPSERSDKGIYTLKLENRVKTTISGEIDVNVVIARPSAPKEL 6997
QY 133 ----- 132
Db 6998 KFGDITKDSVHLTWBPPDDGGSPLTGYVVEKREVSRTWTWKVMDPVTDLFTVPLVQG 7057
QY 133 ----- 132
Db 7058 KEYLFKVCARNKCGPEPAYVDEPNMSTPATVPDPENVKWRDRTRANSIFLTWDPKND 7117
QY 133 ----- 132
Db 7118 GGSRIKGYIVERCPRGSKWACGEPVAETKMVEVTGLEBKWYAYRVKALNRQGASKPSR 7177
QY 133 ----- 132
Db 7178 PTEEIQAVDTQEAPEIFLDVLLAGLVKAGTKIELPATVTKPEPKITWTKADMILKQD 7237
QY 133 ----- 132
Db 7238 KRITIENVPKSTVTIVDSKRSDTGYIIIEAVNCGRATAVVEVNVLDKPGPPAAFDITD 7297
QY 133 ----- 132
Db 7298 VTNESCLLTWNPDRDDGSKITNYVVERRATSEVWHKLSSTVKDTNFKATKLIPNKEYI 7357
QY 133 ----- 132
Db 7358 FRVAENMVGEPVOASPIAKYQFDPGPPTRLEPSDITKDAVTLTWCEPDDGGSP 7417
QY 133 ----- 132
Db 7418 TGYWYERLDPDTKWVRNKNMPVKDTYRVKGLTNKKYRFRVLAENLAGPGKPSKSTEP 7477
QY 133 ----- 132
Db 7478 ILIKDPIDPPPGKPTVKDVGKTSVRLNWTXPEHGGAKIESYVIEMLKTKTGTDEWVRVA 7537
QY 133 ----- 132
Db 7538 EGVPTTQHLPLGLMEGQESYFRVRAVNKAGESEPSPDVLCKREKLYPPSPRWLEVIN 7597
QY 133 ----- 132
Db 7598 ITKNTADLKWTVPKDGSGSPITNYIVEKRDVRRKGQTVDTTIVKDTCTVTPLTEGSLYV 7657
QY 133 ----- 132
Db 7658 FRVAENALGOSDYTEIEDSVLAKDTFTTPGPPYALAVVDVTKRHHVLDLKWEPKNDGGRR 7717
QY 133 ----- 132
Db 7718 IQRYVIEKKERGLTRWVAKGTAGPCNFRVTDVIEGTEVQFQVRAENAGVGHPSSEPT 7777
QY 133 ----- 132
Db 7778 ILSIEDPTSPSPPLDLHVTDAGRKHIAIAWKPPBKNGSGSPIIGYHVMCPVGTCKMWRV 7837

Qy	133	-----	132
Db	7838	NSRPIKDLKFKEVGVPDKEYLVRVRAVNAIGVSEPSEISENVVAKDPCKPTIDLETH	7897
Qy	133	-----	132
Db	7898	DIIVIEGEKLSIPVPRFRAVPVTVSHHKDGKEVKASDLTMKNDHISAHLEVPKSVRADA	7957
Qy	133	-----	132
Db	7958	GIYTTITLENKLGASATASINVKVIGLPGCKDIKASDITKSSCKLTWEPPFDGGTPIIHY	8017
Qy	133	-----	132
Db	8018	VLERREAGRTYIPVMSGENKLSWTVKOLIPNGEYFFRVKAVNKVGGGEYIELKNPVIQA	8077
Qy	133	-----	132
Db	8078	DPQOPDPDPPDVEVHNPTAEAMTITWKPLYDGGSKIMGYIIEKIAKEERWKCNEHLV	8137
Qy	133	-----	132
Db	8138	PILTYTAKGLEGEKEYQFRVRAENAAAGISEPSRATPTTKAVDPIDAPKVKILRTSLEVKRG	8197
Qy	133	-----	132
Db	8198	DEIALDASISGSPYPTITWIKDENVIVPEIKKRAAPLVRRRKGEVOEEFFVLPLTQRL	8257
Qy	133	-----	132
Db	8258	SIDNSKKGESQLVRDSLPRDHGLYMIKVENDHGIKAPCTVSLDTPGPPINVFEDIR	8317
Qy	133	-----	132
Db	8318	KTSVLCKWEPLDDGSEIINYTLEKKDKTKPDSEMIVVTSTLRCHKYSVTKLIEGREYL	8377
Qy	133	-----	132
Db	8378	FRVRAENRFGPGPCYKPLVAKDPGPPADPKFIVEDVTSNMLYKWNKPEKDNKGSPIL	8437
Qy	133	-----	132
Db	8438	GYWLEKEVNSTHWSRVNKSLLNALKANVDGLLEGLTVFVRVCAENAAAGCKFSPPSDPK	8497
Qy	133	-----	132
Db	8498	TAHDPIPPGPPIPRIVTDTSTTIELEWEPAPFANGGEIVGYFVDKQLVGTNWSRCTEK	8557
Qy	133	-----	132
Db	8558	MIKRVQYTVKEIREGADYKLRVSAVNAAGEGPPGETQPTVVAEPQEPPAVELDVSXGGI	8617
Qy	133	-----	132
Db	8618	QIMAGKTLRIPAVVTGRPVPTKWTYKEGELDKDRVIDNVGTSELIIDALKRKHGRY	8677
Qy	133	-----	132
Db	8678	VITATNSCGSKFAARVEYFDVPGPVLDLKPVVTNRKMCLLNWSDPEDDGGSEITGFIIE	8737
Qy	133	-----	132
Db	8738	RKDAKHTWRQPIETERSKCDITGLLEGQEKYFRVIAKNKFCGPPVEIGPILAVDPLGP	8797
Qy	133	-----	132
Db	8798	PTSPERLTYTERTKSTITLDWKEPRNGGSPIQGYIIEKRRHDKPDFERVNKRLCPTTSF	8857
Qy	133	-----	132
Db	8858	LVENLDEHQMYFRVKAVNIGESEPSPLNVVIOQDEVPTTIKURLSVRGDTIKVKAGE	8917
Qy	133	-----	132

Db	8918	PVHIPADV	TGLPMPK1EWSKNETV1EKPTDALQITKEEVSRSSEAKTELISPKAVREDKGT	8977
Qy	133	-----	-----	132
Db	8978	YTVTAS	NRLGSVFRNVHVEYDRPSPRNLAVTDIKAESCYLTWDAPLDNGGSEJTHVVI	9037
Qy	133	-----	-----	132
Db	9038	DKRDSR	KAEWEVENTAVEKRYGIWKLPNGOYEFVRVAVNKYGISDECKSDKVVIQD	9097
Qy	133	-----	-----	132
Db	9098	PYRLPG	PGPKVLARTKGSMLVSWTPPLDNGGSPITGVWLEKREBGSPIYRSVRAPIT	9157
Qy	133	-----	-----	132
Db	9158	KVGLGV	EFNVPRLLGCVKYQFRAMAINAAGIPPEPSDPEVAGDPIFPGPPSPCEVK	9217
Qy	133	-----	-----	137
Db	9218	DKTKSS	ISLWGKPKPAKDGSGPIKGYIVEMQEGTTDWKRVNBPDKLITTCBVCVPLNEL	9277
Qy	138	-----	-----	137
Db	9278	RKYFRV	KAVNEAGESEPSDTTGEIPATDIOBEPEVFDIGAQCCLVKAGSQIRIPAVI	9337
Qy	138	-----	-----	137
Db	9338	KGRTPK	SSNEFDGAKKAMKOGVHDIPEDAQLETAENSSVIIPECKRSHGKYSITAK	9397
Qy	138	-----	-----	137
Db	9398	NKAGOT	ANCVRKVMVDPGPPKDLKVSITRGSCRLSNKMPDDGGDRIKGYVIEKRTID	9457
Qy	138	-----	-----	137
Db	9458	GKAWTK	VNPDGSTTFVVPDLLSEOOYFFRVRARENRFGIPGPVETIQRTTARDPIYPPDP	9517
Qy	138	-----	-----	137
Db	9518	PIKLKIG	LITKNTVHLSWPKPKNDGGSPVTHYIVECLAWDPCTGCKEAWRQCNRDVEEL	9577
Qy	138	-----	-----	137
Db	9578	QFTVED	LVEGGEYFRVKAVNAAGSKPSATVGPCDCQRPDMPPSIDLKEFMEVEEGTNV	9637
Qy	138	-----	-----	137
Db	9638	NIVAKIK	GVFPPTLTFWKAPPKKPNKEPVLVYDTHVYNKLVDVDTCTLVIPOSRHSDTGLY	9697
Qy	138	-----	-----	137
Db	9698	TITAVNL	GTASKEMRLNVLGRGPPVGPIKPEVSADQMTLSWPPPKDGGSKITNVVI	9757
Qy	138	-----	-----	137
Db	9758	EKREANR	KTWVHVSSEPKECTYTIPLKLEHEVVFVFRMAQNKYGIGEPILDSEPTARNLF	9817
Qy	138	-----	-----	137
Db	9818	SVPGAP	KPTVSSVTRNSMTVNNEEPEYDGGSPVTGYWLEMKDTTSKRWKRVNRDPIKAM	9877
Qy	138	-----	-----	137
Db	9878	TLGSYK	VTGLIEGSDYOFVYAINAAGVPASLPSPDAPARDPIAPFPFPKVDWTK	9937
Qy	138	-----	-----	137
Db	9938	SSADLEW	SPPIKGGSKVTGYIVEYKEBGEKEWKGDKVEVRGTKLVTGLKEGAFYKFR	9997
Qy	138	-----	-----	137

```

Db 9998 VSAVNIAGIGEVEYTDVIEMKDLVSPDLQLDASVRDRIYVHAGVIRIIAYSGKPPP 10057
QY 138 ----- 137
Db 10058 TVTWMNERTLPOEATIIETATISSVMVINCORSHQGVYSLAKNEAGERKKTIIIVDL 10117
QY 138 ----- 137
Db 10118 VPGPVGTPFLAHLNLTNESCCLTWFSPEDDGGSPITNYVIEKRESRRAMTPVTYVTRQN 10177
QY 138 ----- ELTKNOVS 145
Db 10178 ATVGLIOQKAVFFRIAAENSGMPFVETSEALVIREPIVPERPEDELVKEVTKNVT 10237
QY 146 LT----- 147
Db 10238 LTWNPPKYDGGSEIINYLESRLIGTERFHKVTNDNLLSRKYTVKLGEGDTYEVRSVAV 10297
QY 148 ----- 147
Db 10298 NIVGOKPSFCTKPICTCKDELAPTLHLDFRDKLTIIRVGEAFALTGRYSRGPVKVSWFK 10357
QY 148 ----- CLV-----KGF----- 153
Db 10358 DEADVEDDRTHIKITPATLALEKIKAKRSDSGKYCVVVENSTGSRKGFQCVNVVDRGP 10417
QY 154 ----- 153
Db 10418 PVGPVSEVTKDYMVISWKPPLDGGSKITNYIIERKEVGKDVMPVTSASAKTTCKVS 10477
QY 154 ----- 153
Db 10478 KLEBGDYIFRIHAENLYGSDPLVSDSMKAKDRFRVPDAPQPIVTEVTKDSALVTWVK 10537
QY 154 ----- 153
Db 10538 PHDGGKPTINYILEKRETMKRWARVKDPIHPYTKFRVPDLLBGCOYEFVRSANEIGI 10597
QY 154 ----- 153
Db 10598 GDSPPSKPVFAKDPIAKSPVPVNEAIDTTCNSVDLTWQPRHDGSKILGYIYEQV 10657
QY 154 ----- 153
Db 10658 GDEWRANHTPESPETKYVTGLRDQYKFRVLAVNAGESDPAHVPEPVLVDRLE 10717
QY 154 ----- 153
Db 10718 PPELILDANMAREQHIKVGDTLRLSAIIGVFPFKVTWKEDRDAPTARIDVTPVGSKL 10777
QY 154 ----- 153
Db 10778 EIRNAHEDGGIYSLTVENPAGSKTVSVKVLVDKPGPPRDLVSEIRKDSCLTWKEPL 10837
QY 154 ----- 153
Db 10838 DDGGSVITNVYVRRDVAQAOWSPLSATSKKSHFAKHLNEGQYLFVRAENQYGRGPF 10897
QY 154 ----- 153
Db 10898 VETPKIKALDPLHPPGPKDLHHVDVKTEVSLVWVKPDRDGGSPITGYLVEYQEGTQ 10957
QY 154 ----- 153
Db 10958 DWIKFXTVNLCEVVTGLOQKTYRFRVKAENIVGLGLPDTTPIECQKLVPPSPVELDV 11017
QY 154 ----- 153
Db 11018 KLIEGLVWAGTVRPPAIIRGVPVPTAKWTTDGSSEIKTDEHYTVETDNFSSVLTIKNCL 11077
QY 154 ----- 153
Db 11078 RDTGEYQITVSNAGSKTVAVHLTVLDVPGPPTGPIINILDTVPEHMTISWQPPKDDGS 11137

```

```

QY 154 ----- 153
Db 11138 PVINYIVEKQDTRKDTWGVVSSGSSKTKLIPHLOKGEYFVRVAENKIGVPPPLDSTP 11197
QY 154 ----- 153
Db 11198 TVAKHKFSPSPGPKVPVTDITENAAVSWTLPKSDGGSPITGYMYMERREVTKGWRVVK 11257
QY 154 ----- 153
Db 11258 TPIADLKFRVTGLVEGNTYEFVFAENLAGLSKSPSSDPIKACRPIKPPGPPINPKLKD 11317
QY 154 ----- 153
Db 11318 KSRETADLVWTKPLSDGGSPILGYVVECOQKPGTAQWNRINKDELIRQCAFVPGLIEGNE 11377
QY 154 ----- 153
Db 11378 YRFRIKAANIVGEGEPRELAESVIAKDIILHPPEVELDVTCRDVITVRVGQTIRILARVKG 11437
QY 154 ----- 153
Db 11438 RPEPDITWTKREGKVLVREKRVLDLIQDLPRVELOIKEAVRADHGKYIISAKNSGHAQSA 11497
QY 154 ----- 153
Db 11498 IVNVLDPRGPCONLKVINTVKENCTISWENPLDNGGSEITNFIVEYRKPNOKGHSIVASD 11557
QY 154 ----- 153
Db 11558 VTRKLIKANLANNEYFVCAENKVGVPETIETKPTILAINPIDRPEPENLHIADGK 11617
QY 154 ----- 153
Db 11618 TFVYLKWRPDPYDGGSPNLSTYHVERRLUKGSDDWERVHKSGIKETHYMWDRCVENOIYEFR 11677
QY 154 ----- 153
Db 11678 VQTKNEGGSWVKTEBWWVVKEDLQKPVLDLKLGLVLTVRAGDTIRLEAGVRGKPFPEVA 11737
QY 154 ----- 153
Db 11738 WTKDKDATLTRSPRVKIDTRADSKFSLTAKAKRSDGGKYVVVATATAGSFVATVNVL 11797
QY 154 ----- 153
Db 11798 DKPGPVRLKIVDVSSDRCTVCHDPPEDDGGCEIQNYILEKCETKRMVWSTYSATVLTGP 11857
QY 154 ----- 153
Db 11858 TTVTRLIEGNEYIFRVRAENKIGTPPTESKPVIAKTKYDKPGRPDPPEVTKYSKBMTV 11917
QY 154 ----- 153
Db 11918 VWNPPYDGGKSTGYFLEKKEKHSTRVWPVNSAISIPERRMKYONLLPDHEYQFVKAEN 11977
QY 154 ----- 153
Db 11978 EIGIGESLPSRPVAKDPIEPPGPTNFRVVDTTKHSITLGHGKPYDGGAPIIGYVVE 12037
QY 154 ----- 153
Db 12038 MRPKIADASPDEGWKRCNAAQLVRKEFTVTSLEDNEQEYEFVCAQNOVGIGRPAELKEA 12097
QY 154 ----- 153
Db 12098 IKPKEILEPPEIDLDASMRKLVIVRAGCPIRLFAIVRGRPAKVTWRKVGIDNVVRKGQV 12157
QY 154 ----- 153
Db 12158 DLVDVTMAFLVIPNSTRDDSGKYSLTILVNPAGEKAFVFNVRVLDTPGVSCLKVSDVTKTS 12217

```

QY 154 ----- 153
Db 12218 CHVSWAPPENDGGSQVTHYIVKREADRTWSTVTPVKKTSFHTNLNPGNEYFRVTA 12277
QY 154 ----- 153
Db 12278 VNEYGPGVPTDVKPVLASPLSEDPDRKLEVTEMKNSATLAWLPPLRGGAKIDGYI 12337
QY 154 ----- 153
Db 12338 TSYREEQPADRWTEYSVVKDLSLVITGLKEGKKYKFRVAARNAGVSLPREAEGVYEAK 12397
QY 154 ----- 153
Db 12398 EQLPPRIILMEQITIKAGKRLRIEAHVYKPHPTCKWKKEDEVVTSSHLAVHKADSS 12457
QY 154 ----- 153
Db 12458 ILIKDVRKDSGYSLTAENSGTDQKIKVVVMDAPGPPQPPFDISDIDADACSLSWH 12517
QY 154 ----- 153
Db 12518 IPLEDGGSNITNYIVEKCDVSRGDWVTALASVTKTSCRVKLLIPGOEYIFRVAENRFGI 12577
QY 154 ----- 153
Db 12578 SEPLTSPKMAQPPFGVPSEPKNARVTKVNKDCIFVANDRPDSGGSPILIGYLIERKERN 12637
QY 154 ----- 153
Db 12638 SLLWVKANDTLVRSTEYPCAGLVEGLEYSFRIYALNKAGSPSPKTEYVTARMPVDPG 12697
QY 154 ----- 153
Db 12698 KPEVIDVTKSTVSLIWARPKHDGSKLIIGYFVEACKLPGDKWVRNCTAPHQIQEEYAT 12757
QY 154 ----- 153
Db 12758 GLEKAOYQFRAIARTAVNISPPSEDPVTILAENVPPRIDLSVAMKSLITVKAGTNVC 12817
QY 154 ----- 153
Db 12818 LDATVFGKPMPTVSWKDGTLKPAEGIKMAMORNCTLELFSVNRKDSGDTITAENSS 12877
QY 154 ----- 153
Db 12878 GSKSATIKLVLDKPGPPASVKINKMYSDRAMLSWEPPLEDGGESEITNYIVDKRETSRPN 12937
QY 154 ----- 153
Db 12938 WAQVSATVPITSCSVEXLEIEGHEYQFRIACENKYGVGDPVFTPEAIKPNFYDPPGRCDPP 12997
QY 154 ----- 153
Db 12998 VISNITKDHMTVSWKPPADGGSPITGYLLEKRETOAVNWTKVRKPIIERTLKATGLQE 13057
QY 154 ----- 153
Db 13058 GTEYFRVTAINKAGPCKPSDASKAAYARDPQPPAPPAPKPVYDTRSSVLSWGPAY 13117
QY 154 ----- 153
Db 13118 DGGSPILIGLVEVKRADSNWVRNLPQNLQKTRFVYGLMEDTQYQFRVAVNKGYS 13177
QY 154 ----- 160
|| || :
Db 13178 PSDVPDRKHPKDLILIPPEGLADLRKTLILRAGVTMRLYVPVKGRPPPKITWSKPNVL 13237
QY 161 ----- 160
Db 13238 RDRIGLDIKSTDDFTFLRCENVNKYDAGKYILILENSCGKKEYTIVVKVLDTPGPPNVNT 13297
QY 161 ----- 160

Db 13298 VKEISDSAYVTWEPPIIDGGSPIIINYVQKRDARERSWSTVTTECSKTFRVANLEEGK 13357
QY 161 ----- 160
Db 13358 SYFERFAENEYIGIDPGETRDVAKASQTPGPVVDLKVRSVSKSSCSIGWKKPHSDGGS 13417
QY 161 ----- 160
Db 13418 IIGVVVDFLTENKWRVMKSLSLQYSAKDLTEGKEYTFRVSAENEGEPTSEITVVAR 13477
QY 161 ----- 160
Db 13478 DDVAPDLDLKGLPDLCLYLAKENSFRLKIPKGPAPSVWKKGEDPLATDTRVSSESS 13537
QY 161 ----- 160
Db 13538 AVNTLLIVDCQSDAGKYITILKNVAGTKEGTISIKVVGKPGIPTGPIKFEVTAEMT 13597
QY 161 ----- 160
Db 13598 LKWAPPKDDGSEITNILEKRDSVNKNKWTCASAVQKTTFRVTRLHEGMEYTFRVSAN 13657
QY 161 ----- 160
Db 13658 KYGVGGLKSEPIVARHPFDVDPAPPPNIVDVHRDSVSLTWTDPKKTGGSPITGYHLEF 13717
QY 161 ----- 160
Db 13718 KERNSLLWKANKTPIRMROFKVTGLTEGLEYEFRVMAINLAGVKFSLPSEPVALDPI 13777
QY 161 ----- 160
Db 13778 DPCKPEVINITRNSVTLIWTEPKYDGHKLTGYIVEKRODLPKSKWKNHVNYPECAFT 13837
QY 161 ----- 160
Db 13838 VTDLVEGGKEYFRIRAKNTAGAISAPSESTETIICKDEYAPTIVLDPTIKDGLTIRAGD 13897
QY 161 ----- 160
Db 13898 TIVLNAISILGKPLPKSSWSKAGKDIRPSDITQITSTPTSMILIKYATRKDAGEYITA 13957
QY 161 ----- 160
Db 13958 TNPFGTKVEHVKTIVLDVPGPPGPEISNVSAEKATLTWTPTLEDGGSPIKSYILEKRET 14017
QY 161 ----- 160
Db 14018 SRLLTWTVSEDIQSCRHVATKLIQONEYIFRVSANHYGKGEYQSEPVKMDRFGPPGP 14077
QY 161 ----- 160
Db 14078 PEKPEVSNVTKNTATVSWKRPVDDGGSEITGYHVERREKKSRLVRAIKTPVSDLRCKVT 14137
QY 161 ----- 160
Db 14138 GLOGSTYFRVSAENRAGIGPSEASDSVLMKDAAYPPGPPSPNPHVTDITTKKSASLAWG 14197
QY 161 ----- 160
Db 14198 KPHYDGGLEITGYVVEHQVKGDEAWIKDTTGTLRITOFVVVPDLQTKRYNFRISAINDA 14257
QY 161 ----- 160
Db 14258 GVCEPAVIPDVEIVEREMAPDFELDAELRRLTVVRAGLSIRIFVPIKGRPAPEVTWKDN 14317
QY 161 ----- 160
Db 14318 INLKNRANIENSTFTLLIIPECNRYDTGKFVWTIENPAGKSGFVNVRVLDTPGVLNL 14377
QY 161 ----- 160

Db 14378 RPTDITKDSVTLHWDLPDLIDGGSRIITNYIVERKREATRKSISTATTKCHCKTYKVITGLSEG 14437
QY 161 ----- 160
Db 14438 CEYFFRMAENEYIGCEPTETPEPVKASEAPSPDLSNIMDITKSTVSLAWPKPKHDGGS 14497
QY 161 ----- 160
Db 14498 KITGYVIEAQRKSDQWHTITTVKGLECVVRNLTEGEYTFQVMVNSAGRSAPRESRPV 14557
QY 161 ----- 160
Db 14558 IVKEQTMPELDLRGIYQKLVIAKAGDNKIVEIPVLRPKPTVWKKGDIILKOTORVNF 14617
QY 161 ----- 160
Db 14618 ETTATSTILNINECVSDSGPYPLTARNIVGEVDVITIQVHDIPGPTGPIKFDVSSD 14677
QY 161 ----- 160
Db 14678 FVTFSDPPDNGGVPISNVVEMRQDSTTWVELATTVIRTTYKATRLTTGLEYOFRVK 14737
QY 161 ----- 160
Db 14738 AQNRVGVGPIITSACIVANYPFKVPGPGTPQVAVTKDSMTISMHEPLSDGGSPIILGYH 14797
QY 161 ----- 160
Db 14798 VERKERGILWQTVSKALVPGNIFKSSGLTDGIAEFVRVIAENMAGKSKPSKPSBPLAL 14857
QY 161 ----- 160
Db 14858 DPIDPPKVPPLNITRHTVTLKWAPEYTGFKITSYIVEKRDLPNGRWLKNFNSILEN 14917
QY 161 ----- 160
Db 14918 EFTVSLGTEDAAEFVRVIAKNAAGAISSPPSEPSDAITCRDDVEAPKIKVDVRFKDTVILK 14977
QY 161 ----- 167
Db 14978 AGEAFRLADVSRPPPTMEWSKDGKLEGTAKLEIKIADFSTNLVKNKDSIRDSGAYTL 15037
QY 168 ----- 167
Db 15038 TATNPGGFAKHIFNVKVLDRPGPEGLAVTEVTSEKVLFWFPLDDGGAKIDHYIVOK 15097
QY 168 ----- 167
Db 15098 RETSLRANTVASEVOVTKLKTKLLKGNEYIFRVMVNNKYGVGEPLSEPEVLAVNPYGP 15157
QY 168 ----- 167
Db 15158 PDPPKNPEVTTITKDSMVVWGHGPDSDGSEIINYIVERRRKAGORWIKCNKKTLDLRY 15217
QY 168 ----- 167
Db 15218 KVSGLTEGHEYEFRIEMAENAGISAPSPSPFYKACDTVFKPGPPGNPRVLDTSRSSISI 15277
QY 168 ----- 167
Db 15278 ANKPIYDGGSEITGYMVEIALPEDEWOIVTPPAGLKATSYITITGLTENQYKIRIYAM 15337
QY 168 ----- 167
Db 15338 NSEGLGEPALVPGTPKAEDRMLPPELELDADLRKVVTIRACCTLRFLFPIKGRPAPEVKW 15397
QY 168 ----- 167
Db 15398 ARDHGESLKDASTESTSSYTLIVGNVRNFRDVGKVLITVENSNGSKSAFVNVRLDTPG 15457
QY 168 ----- 167
Db 15458 PODLKVKEVTKTSVLTWDPPLDGGSKIKNYIVEKRESTRKAYSTVATNCHKTSWKVDQ 15517

QY 168 ----- 167
Db 15518 LQEGCSYFPRVLAENEYIGLPAETAESVKASERPLPPCKITILMDVTRNSVLSWEKPEH 15577
QY 168 ----- 167
Db 15578 DGGSRILGYIVEMQTKGSDKWATCATVKVTEATITGLIOGEEYSFRVSAQNEKGISDPHQ 15637
QY 168 ----- 167
Db 15638 LSVPVIAKDLVIPPAPKLLFNTFTVLAGEDLKVDVPFIFGRPTPAVTHKDNVPLKQTTVR 15697
QY 168 ----- 167
Db 15698 NAESTENNSLLTIKDACREDVGHYVVKLTNSAGEAETLNVLDPKPGPTGPVKMDEVY 15757
QY 168 ----- 167
Db 15758 ADSITLSWGPYPKYDGGSSINNYIVEKRDJSTTTTQWIVSATVARTTIKACRLKTGCEYQFR 15817
QY 168 ----- 167
Db 15818 IAAENRYGKSTYLNSEPTVAQYPFVKVPGPGTPVVTLSRDSRDSMEVQWNEPISDGGSRVIG 15877
QY 168 ----- 167
Db 15878 YHLERKERNISILWVKNKTPIPQTKFTGTLEEGVEYEFVSAENIVGIGKPSKYSECV 15937
QY 168 ----- 167
Db 15938 ARDPCDPGRPEAIIVTRNSVTLOWKKPTYDGGSKITGYIVEKKELPEGRMWSKFTNII 15997
QY 168 ----- 167
Db 15998 DPHFEVTLVEDHRYEFVRVIAARNAAGVSEPSSESTGAIARDEVDPPIRISMDPKYKDTIV 16057
QY 168 ----- 167
Db 16058 VHAGESFKVDADIYKPIPTIQWIKGDOELSNARLEIKSTDFATSLSVKDAVRVDSGNY 16117
QY 168 ----- 167
Db 16118 ILKAKNVAGERSVTVMKVLDPRPGPEGVVIVSGVTAEKCTLAWKPPLODGGDIINYIV 16177
QY 168 ----- 167
Db 16178 ERRETSRLVWTVDANVQTLCKVTKLLEGNEYTFRIMAVNNKYGVGEPLSEPEPVAKNPF 16237
QY 168 ----- 167
Db 16238 VVPDAPKAPVTTVTKDSMIVVWERPASDGGSEILGYVLEKKRDEGIRWTRCHKRLIGEL 16297
QY 168 ----- 167
Db 16298 RLVRTGLIENHDYEFVSAENAGLSEPPSPAYQKACDPIYKPGPPNNPKVIDITRSSV 16357
QY 168 ----- 167
Db 16358 FLWSKPIYDGGCEITQGYIVEKCDVSGEWTWCTPPTGINKTNIEVEKLEKHEYNFRIC 16417
QY 168 ----- 167
Db 16418 AINKAGVEHADVPPIIVEEKELEAPDIDLLELRKIINIRAGGSLRULFVPIKGRPTBV 16477
QY 168 ----- 167
Db 16478 KWGVDEIRDAAIIDVTSSFTSLVDNWNRYDSGKTLTLENSGTSKSAFVTVRLDTP 16537
QY 168 ----- 167
Db 16538 SPPVNLKVTEITKDSVSTWBPPLDGGSKIKNYIVEKREATRKSAAVAVTNNCHNSWKI 16597

QY 168 ----- 167
Db 16598 DQLEGCXYFRVTAENEYIGLPAQTADPIKVAEVPQPGKITVDDVTRNSVLSWTKP 16657
QY 168 ----- 167
Db 16658 EHDGSKIIOYIVEMOAKHSEKSECARVKSQAQVITNLTOGSEYLFRRVAVNEKGRSDP 16717
QY 168 ----- 167
Db 16718 RSLAVPIAKDLVIEDVRPAFSSYSVOVQODLKIEVPISGRPKPTITWTCKDGLPLKQTT 16777
QY 168 ----- 167
Db 16778 RINWTDLDLTLISIKETHKDDGGQYITVANVVGQKTASIEIVITLDKDPKPGVKFDD 16837
QY 168 ----- 167
Db 16838 VSAESITLSNPLYTGGCOITNYIVQKRDTTTVWDVVSATVARTLKVTKLKTGEYO 16897
QY 168 ----- 167
Db 16898 FRIFAENRYCOSFALESDPITVAQYPIKEPGPGTPTAFALSKDSMVIOHHEPVNNGSPV 16957
QY 168 ----- 167
Db 16958 IGYHLERKERNILWTKVNTKIIHDTQFAQNLEEGIEYEFVRYAENIVGVGKASKNSEC 17017
QY 168 ----- 167
Db 17018 YVARDPCDPGTPPEPIMVRNEITLQWTKPVYDGGSMITGYIVEKRDLPDGRWMAKSFN 17077
QY 168 ----- 167
Db 17078 VIETQTVSGLTEDQRYEFVRIAKNAAGALSKPDSSTGPTAKDEVELPRISMDPKFRDT 17137
QY 168 ----- 167
Db 17138 IVNAGETFRLEADVHGKPLPTIEMLRGDKEIEBSARCEIKNTDFKALLIVKDAIRIDGG 17197
QY 168 ----- 167
Db 17198 QYILRASNAVAGSFPVNVKVLDRPDPGPEGVQVGTSEKSLTWSPLQDGGSDISHY 17257
QY 168 ----- 167
Db 17258 VVEKRETSRLAWTVVASEVVTNSLKVTKLLEGNEYVFRIMAVNKYGVGEPLSAPVLMKN 17317
QY 168 ----- 167
Db 17318 PFVLPGPKSLEVTNIAKDSMTVCWNRPSDGGSEIIGYIVEKRDGRGIRWIKCNKRIT 17377
QY 168 ----- 167
Db 17378 DLRLVTLGTEDHEYEFVRSNAENAGVGEPSATVYKACDPVKPGPPTNAHIVDTKN 17437
QY 168 ----- 167
Db 17438 SITLANGKPIYDGSSEILGVYVEICKADEEWQIVTPTQCLRVTRPEISKLETHEQYKIR 17497
QY 168 ----- 167
Db 17498 VCAINKVGLGEATVSPGTVPKPEDKLEAPELDLDSELKRGIVVRAGGSARIHIPFKGRPTP 17557
QY 168 ----- 167
Db 17558 EITWSREGEFTDKVOIEKGVNTQLSIDNCDRNDAGKYILKLENSGSKSAFVTVKVLVD 17617
QY 168 ----- 170
Db 17618 TPQPPQNLAVKEYRKDSAFVLEWPEPIIDGGAKVKNYVIDKRESTRKAYANVSSKCSKTSF 17677
QY 171 ----- 170

Db 17678 KVENLJEGAIYFRVMAENEFVGVVETVDVAKAAEPPSPGKVTLTLDVSQTSASLWE 17737
QY 171 ----- 170
Db 17738 KPEHGGSRVLGVVEMQPKGTEKWSIVAESKVCNAVVTGLSSGOEYQFRVKAYNEKGKS 17797
QY 171 ----- 170
Db 17798 DPRVLGVPVIAKDLTIQPSLKLFPNTYSIOAGEDLKEIPVIGRPRNISMVWKGDEPLKQ 17857
QY 171 ----- 170
Db 17858 TTRVNVEETATSTVLHIKEGNKDDFGKYTVTATNSAGTATENLSVIVLEKPGVPVGRF 17917
QY 171 ----- 170
Db 17918 DEVSADFVVISWEPAYTGGCQISNYIVEKRDTTTTTMMHVSATVARTTIKTKLTGTE 17977
QY 171 ----- 170
Db 17978 YQFRIFAENRYKSAPLDSKAVIVQYPPKPPGTPFTVISKQDQMLVOWHEPVNDGGT 18037
QY 171 ----- 170
Db 18038 KIIGYHLEQKEKNSILWVKLNKTPIQDTKFKTGLDEGLEEFKVSNAENIVGIGKPSKVS 18097
QY 171 ----- 170
Db 18098 ECFVARPCDPPGRPEAIVITRNNVTLKWKPAYDGGSKITGYIVEKDDLPDGRWMAKSF 18157
QY 171 ----- 170
Db 18158 TNVLETFVTVGLVEDQRYEFVRIARNAAGNFSEPSDSSGAIITARDEIDAPNASLDPKYK 18217
QY 171 ----- 170
Db 18218 DVIWHAGETFVLEADIRGKPIPDVVWVSKDGKELEETAARMEIKSTIQKTTLVVKDCIRT 18277
QY 171 ----- 170
Db 18278 DGGQYILKLSNVGTSIPITVVKLDRPDPGPEGLKVTGYTAEKCYLAWNPLODGGANI 18337
QY 171 ----- 175
Db 18338 SHVILIEKRETSRLSWTQVSTEVOALNKVTKLLPGNEYIFRVMVKNKYGIGEPLESGPVT 18397
QY 176 ----- 175
Db 18398 ACNPYKPPGPPSTPEVSATKDSMVVTWARPVDDGGTEIEGYILEKRDKEGVRTKCNKK 18457
QY 176 ----- 175
Db 18458 TLFDLRLVTVGLTEGHSYEFVRAAENAGVGEPSFVFYRACDALLYPPGPPSNPKVTD 18517
QY 176 ----- 175
Db 18518 SRSSVSLANSKPIYDGGAPVKYGVVEVKEAAADEWTCTPTTGLQKQFTVTKLKENTEX 18577
QY 176 ----- 175
Db 18578 NFRICAINSEGVGEPATLPGSVVAQRIEPPETELDADLRKVVVLRASATLRLFVTIKGR 18637
QY 176 ----- 175
Db 18638 PEPEVKWEKAEGILTDRAQIEVTSSTMLVIDNVTRFDSGRYNLTLENNSGSKTAFVNV 18697
QY 176 ----- 175
Db 18698 VLDSPAPVNLITREVKKDSVTLSEWPEPLIDGGAKITNYIVEKRETRKAYATITNNCTK 18757
QY 176 ----- 175

Db 18758 TTFRIENLQGGCSYFPRVLASNEYGIGLPAETTEPVKVSEPLPGRVTLVDVTRNTATI 18817
QY 176 ----- 175
Db 18818 KWEPESDGGSKITGYVVEWQTKGSEKWCSTCQVKTLEATISGLTAGEEYVFRVAANVK 18877
QY 176 ----- 175
Db 18878 GRSDPRLGVPVIARDIEIKPSVELPFHTFNVKAREQLKIDVPFKGRPOATVNRKRGOT 18937
QY 176 ----- 175
Db 18938 LKETRNVSSSKVTVLSIKEASKEDVGYELCVSNSAGSITVPITIIIVDRPPGPI 18997
QY 176 ----- 175
Db 18998 RIDEVSCDSITISWNPPEYDGGCQISNYIVEKKEITSTTHIVSQAVARTSIKIVRLTIG 19057
QY 176 ----- 175
Db 19058 SEYQFRVCAENRYGKSSYSESSAVVAEYFPSPGPGTPKVVHATKSTMLVTWQVPVNDG 19117
QY 176 ----- 175
Db 19118 GSRVIGYHLEYKERSIILMSKANKILLIADTQMKVSGLDEGLMYEYRVYAENIAGICKSK 19177
QY 176 ----- 175
Db 19178 SCEPVPARPCDPGPOPEVTNITRKSLSLWKSHPHYDGGAKITGYIVERRELPGRWLKC 19237
QY 176 ----- 175
Db 19238 NYTNIQETYFEVTELTEDQRYEFRFARNAAADSVSEPSSTGPIIVKODVEPPVMDVK 19297
QY 176 ----- 175
Db 19298 FRDIVVKGAEVLKINADIAGRPLPVISWAKDGBIEIERARTEIISTDNHLLTVKDCIR 19357
QY 176 ----- 175
Db 19358 RDTGVVLTLLKNVAGTRSVAVNKVLDKPGPAGPLEINGLTAEKCSLSWGRPQEDGGAD 19417
QY 176 ----- 175
Db 19418 IDYIIVEKRETLAWTICEGELQMTSCKVTKLLKGNEYIFRVTGVNKGVGCEPLESAI 19477
QY 176 ----- 175
Db 19478 KALDPFTVPSPPTSLEITSVTKESMTLCWSRPESDGSSEISGYIITEREKNSLRWRVVK 19537
QY 176 ----- 175
Db 19538 KPYVDLRVKSTGLREGCEVEYRYAENAAGLSLPSSETPLRAEDPVFLPSPSPKPIVD 19597
QY 176 ----- 175
Db 19598 SGKTTITIAWRKPLEDGGAPITGYVEYKSDTDWKTISIQLRGTEYTIISGLTTGAEV 19657
QY 176 ----- 175
Db 19658 FRVKSVMKVGASDPSDSDPOIAKEREELPLFDIDSEMRKTLIVKAGASFTMTVPFRGR 19717
QY 176 ----- 175
Db 19718 VPNVLSKRPDTLRTRAYVDTTDSRTLTENANRDSGKYTLTIQNVLSAASLTVVKV 19777
QY 176 ----- 175
Db 19778 LDTGPGPTNITVQDVTKESAVLSWDVPENDGGAPVKNYHIEKREASKKAWSVTNNCR 19837
QY 176 ----- 175
Db 19838 SYKVTNLQEGAIYFPRVGENEFGVIPAETKEGVKITEKPSPEKLGVTISIKDSVSLT 19897

QY 176 ----- 175
Db 19898 WLKPEHDGSRIVHYVVEALEKGQKNWKCAVAKSTHHVVSGLRESEYFFRVAENQAG 19957
QY 176 ----- 175
Db 19958 LSDPRELLLPVLIKQLEPEIDMKNEFPHSTVYVRAGSNLKVDIPIISGKPLPKVTLSRDG 20017
QY 176 ----- 175
Db 20018 VPLKATMRNTEITAENLTINLKESVTADAGRYEITAANSSGTTKAFINIVVLDRRPGPT 20077
QY 176 ----- 175
Db 20078 GPVWISDITEESVTLKWEPPKYDGGSQVTNYILLKRETSTAVMTEVSATVARTMMKVMKL 20137
QY 176 ----- 175
Db 20138 TTGEEYQFRIKAENRFGISDHIDSACVTVKLPYTPPGPSTPWVTVNTRESTITVGMHEPV 20197
QY 176 ----- 175
Db 20198 SNGGSVVGYHLEMKDRNSILMQKANKLVIRTHFKVTTISAGLIYEFVRVYAENAGVCK 20257
QY 176 ----- 175
Db 20258 PSHPSEPVLAIADACEPPRNVRITDISKNSVLSLWQOPAFDGGSKITGYIVERRDLDPGRW 20317
QY 176 ----- 175
Db 20318 TKASFTNVTETOFIISGLTQNSQYEFVRFARNAGVISNPSSEVVGPIITCIDSYGPGVIDL 20377
QY 176 ----- 175
Db 20378 PLEYTEVVKYRAGTSVKLRAGISGKPAPIEWKODKELQTNALVCVNTTOLASILKD 20437
QY 176 ----- 175
Db 20438 ADRLNSGCYELKLRNAGMSASATIRVQILDKPGPGPIEFKTVTAEKITLLARPPADGG 20497
QY 176 ----- 175
Db 20498 GAKITHYIVEKRETSRVVMSVSEHLEPCIITTTKIIKNGEYIFRVRVAVNKGIGEPLS 20557
QY 176 ----- 175
Db 20558 DSVAKNAFVTPPGPIPEVTKITKNSMTVWWSRPIADGGSDISGYFLEKRDKKSGLWPK 20617
QY 176 ----- 175
Db 20618 VLKETIRDTROKVTGLTENSQYQYRVCVNAAGOGFSEPFSEFYKAADPIDPPGPPAKIR 20677
QY 176 ----- 175
Db 20678 IADSTKSSITLWSPVYDGGSAVTGYVVEIRQGEETTVSTRGEVTRTEYVVSNLKP 20737
QY 176 ----- 175
Db 20738 GVNYYFRVSAVNACOGGPIEMNEPVQAKDILEAPEIDLVALRTSVIAKAGEDVOVLIP 20797
QY 176 ----- 175
Db 20798 FKGRPPPTVTRKDEKNLGSADARYSIENTDSSLLTIPOVTRNDTKYILLTIENGVBEPK 20857
QY 176 ----- 175
Db 20858 SSTSVKVLDTPAACQKLQVAKHVRGTVTLWMDPPLIDGGSPIINVIYEKRDATKRTWSV 20917
QY 176 ----- 175
Db 20918 VSHKCSSTSFKLIDLSEKTPFFFRVLAENEIGICEPCTETEPVKAEEVAPAPIRDLMSKMS 20977

QY 176 ----- 175
Db 20978 TKTSVILSWTKPFDGGSVITEYVVERKKGEOQWTHSHAGISKTEIEVSQKBEQSVLEPR 21037
QY 176 ----- 175
Db 21038 VFAKNEGLSDPVTIGTITVKELIITPEVDLSDIPGAQVTRIGHNVHLELPYKGPCKPS 21097
QY 176 ----- 175
Db 21098 ISWLKDLPLKESEFFVRFSKTENKITLSIKNAKHEGKKYTVILDNAVCRIAPITVITL 21157
QY 176 ----- 175
Db 21158 GPSPKPGPIRFDEIKADSVILSWDPEDNGGGEITCYIEKRETSQTNKMKVCSVART 21217
QY 176 ----- 175
Db 21218 TFKVNLVKDAEYQFVRRAENRYGVSQLVSSIIVAKHQPRIPGPGPKPVIYNWSDGMS 21277
QY 176 ----- 175
Db 21278 LTWDAPVYDGGSEVTGFHVEKKERNSTLMQKVNTSPISGREYRATGLVEGLDYQFRVYAE 21337
QY 176 ----- 175
Db 21338 NSAGLSSPSPSKFTLAVSPVDPGTPDYIDVTRETITLKNPPLRDGGSKIYGSIEKR 21397
QY 176 ----- 175
Db 21398 QGNERWVRNFTDVSECQYTVTGLSPGDRYEFRIIARNAVGTISPPSQSGIIMTRDENV 21457
QY 176 PPV----- 178
|||
Db 21458 PPIVEGPEYFDGLIIKSGESLRIKALVQGRPVPRVTFWKDGVGEIEKRMNWEITDVLGST 21517
QY 179 ----- 178
Db 21518 SLEVRDNRHGYTVEAKNASAKAEIKVKVQDTPGKVVGPIRFTNITGEKMTLWMD 21577
QY 179 ----- 178
Db 21578 APLNDGAPTHVIIIEKRETSRLAWALIEDKCEAQSYAIIKLNENYQFVRSAVKNFGV 21637
QY 179 ---LSD- 182
|||
Db 21638 GRPLSDPVVAAIQYTVPDAPGIPEPSNITGNSITLTWARPESDGGSEIQYILERREKK 21697
QY 183 ----- 182
Db 21698 STRWVKVISKRPISETRFRKVTGLTEGNEYEFHYMAENAGVGPASGISRLIKCREPVNPP 21757
QY 183 ----- 182
Db 21758 GPPTVVKVTDTSKTVSLEWSKPVFDGGMEIIGYIEMCKADLGDWHKVNABACVKTRYT 21817
QY 183 ----- 182
Db 21818 VTDLOAGEYKFRVSAINGAGKDSCEVTCTIKAVDRLTAPELDIDANFKQTHVVRAGAS 21877
QY 183 ----- 182
Db 21878 IRLFIAYQGRPTTAVWSKPDNSLSRADIHHTDSFSTLTVCNCRNDAGKYTLTVENNS 21937
QY 183 ----- 182
Db 21938 GSKSITFTVKVLDTPGPGPITFKDVTGRSATLMDAPLLDGGARITHYVVERKREARRS 21997
QY 183 ----- 182
Db 21998 WQVISEKCTRQIFKVNDLAEQVYVYFRVSAVNEYGVEPEPIVATEQAPPRRLDV 22057
QY 183 ----- 182

Db 22058 VDTSSAVLAWLKPDDHGGSRITGYLLEMRQKGSDFWEAGHTKQLTFTVERLVEKTEY 22117
QY 183 ----- 182
Db 22118 EFRVKAKNDAGYSEPREAFSSVIIKEPQIEPTADLTGITNOLITCKAGSPFTIDVPISGR 22177
QY 183 ----- 187
-----GSFFL-----
Db 22178 PAKVTVWKLKEMRLKETDRYSIITTKDRTTLTVKDSMRGDSGRYFLTLENTAGVKTFSVT 22237
QY 188 ----- 187

Db 22238 VVIGRPGVPTGPIEVSSVAESCVLSWGEPKDGGTEITNYIVEKRESGTTAQLVNSS 22297
QY 188 ----- 187
Db 22298 VKRTQIKVTHLTXYMEYSFRVSSNREFGVSKPLESAPIIAEHFPVPPSAPTRPEVYHVSA 22357
QY 188 ----- 187
Db 22358 NAMSIRWEEPYHDGGSKIIGYVWEKKERNITLWVKENKVPCLCNKVKTVGLVEGLEVQFR 22417
QY 188 ----- 187
Db 22418 TYALNAGVSKASEASRPIMAQNPVDAPGRPEVTDVTRSTVSLIWSAPAYDGGSKVVGVI 22477
QY 188 ----- 187
Db 22478 IERKPVSEVDGRWLKCNVTIVSDNFTVTALSEGDTYEFRLVAKNAAGVISKSESTGP 22537
QY 188 ----- 187
Db 22538 VTCRDEYAPKAELDARLHGLDVTIRAGDLVLDAAVGGKPEPKI IWTGKDELDCVK 22597
QY 188 ----- 187
Db 22598 SLOYTKRATAVIFKCDRSDSGKYTLTVKNASGTVKAVSMVKVLDSPGCKLTVSRVTQ 22657
QY 188 ----- 187
Db 22658 ECKTLAWSLPQEDGGAETHYIVERRETSRLNWVIVEGECTLSYVVVTRLKKNNEYIFRV 22717
QY 188 ----- 187
Db 22718 RAVNKYGPVVESEPIVARNSTFIPSPGIPPEVGTGKEHI IIQWTKPESDGGNEISNY 22777
QY 188 ----- 187
Db 22778 LVDKREKSLRWTRVNNKYVYDTRLKVTSLMEGCDYQFRVAVNAAGNSEPSEASNFIS 22837
QY 188 ----- 187
Db 22838 CREPSYTPGPPSAPRVVDTTKHSISLAWTKPMYDGGTDIVGYVLEMQEKTDQWYRVHTN 22897
QY 188 ----- 187
Db 22898 ATIRNTEFTVPLDKMGOKYSFRVAANVVKGMSEYSEIAETEPVERIEIPDLELADLKK 22957
QY 188 ----- 195
-----YSKLTVDK-----

Db 22958 TVTIRAGASRLMVSVSGRPPPVITWSKQIDLASRAIIDTTESYSLLIYDKVNRNDAGK 23017
QY 196 ----- 195
Db 23018 YTIEAENQSGKKSATVLVKVYDTPGCPSPVKVEVSRDSVTITWEITIDGGAPVNNYIV 23077
QY 196 ----- 195
Db 23078 EKREAAAMRAFKTVTKCSKTLIRISGLVEGTYMYFRVLPENIYGIGPECSTDAVLVSEV 23137
QY 196 ----- 195

Db 23138 PLVPAKLEVVVDVTKSTVTLANEKPLYDGGSLRTGYVLEACKAGTERMMKVVTLPKPTVLEH 23197
QY 196 ----- 195
Db 23198 TVTSLNEGQYLFRIRAEKGVSEPREIVTAVTQDLRLVTLIDLSTMPQKTIHVPAGR 23257
QY 196 ----- 195
Db 23258 PVELVIAGRPAPPAWFFAGSKLRESERVTVETHKVAKLTIRETITRDTEVTLELK 23317
QY 196 ----- 195
Db 23318 NVTGTTSETIKVILDKPGPTGPIKIDEIDATISISWPELPGAPLSGYVVEQRDA 23377
QY 196 ----- 195
Db 23378 HRPGLPVSESVTRSTFKETRLTEGNEYVFRVAANTNRFGISYLOSEVIECRSSIRIPG 23437
QY 196 ----- 195
Db 23438 PETLIQFVSDGMLTWYPEDDGSQVTGYIVERKEVRADRWVRVNVKVPVTMTYRST 23497
QY 196 ----- 195
Db 23498 GLTEGLEVHRVTAINARGSGKPSRSPKPIVAMPDPIAPPKQPNRPVTDTRTSVSLAWS 23557
QY 196 ----- 195
Db 23558 VPDEGGSKVTGYLIEMQKVDOHEWTKNTPTKIREYTLTHLPOGAEYRFRVLACNAGG 23617
QY 196 ----- 201
Db 23618 PGPAEVPCTVKVTEMLEYPDYELDERYQEGIFVRQGGVIRLTIPIKGRPFPICKWTKREG 23677
QY 202 ----- 201
Db 23678 QDISKRATATSETHVELVIKEADRDGSGTYDLVLENKCGKAVYIKVRIGSPNSPEGP 23737
QY 202 ----- 201
Db 23738 LEYDDIQRVSRVSRPPADGGADILGYILERREVPRKAAWYIDSVRVGTSLVVKGLKE 23797
QY 202 ----- 201
Db 23798 NVEYHFRVSAENQFGISKPLKSEEPVTKPLNPPPEPNPEVLDVTKSSVLSWSRPK 23857
QY 202 ----- 201
Db 23858 DDGSRVTGYIYTERKETSTDKWVRHNKQITTTMYTVTGLVPDAEYQFRIIAQNDVGLSE 23917
QY 202 ----- 201
Db 23918 TSPASEPVVCKDPDKPSQPGLEILSISKDSVTLOWEKPEDCGKEILGYWVEYRQSGD 23977
QY 202 ----- 201
Db 23978 SAWKSKNERIKDKOFTIGGLLEATEYEFVRPAENETGLSRPRRTAMSIKTLTSGEAPG 24037
QY 202 ----- 201
Db 24038 IRKENKDVTKLGEAAQLSCQIVGRPLPDIKWYRFGKELIOSRKYKMSSDGRTHTLTVM 24097
QY 202 ----- 201
Db 24098 EEQEDEGVVTCIATNEVEGETSSKLLQATPQPHPGYPLKEKYGAVGSTLRLHVMYIG 24157
QY 202 ----- 204
Db 24158 RVPAMTWFGQKLLQNSENITIENTHYTHLVKMNQVQKTHAGKYKVQLSNVFGTVDAI 24217
QY 205 ----- 204
Db 24218 LOVEIQDKPDKPTGPVIVIEALLKNSAVISWPKPPADGGSGSWITNYVVEKKEAEGEWOLV 24277

QY 205 ----- 204
Db 24278 SSAISVTTTCRIVNLTENAGYFRVSAQNTFGISDPLEVSSVVIIKSPREKPGAPGKPTIT 24337
QY 205 ----- 208
Db 24338 AVTKDSCVYAMKPPASDGGAKIRNYYLEKREKKQNKWISVTTEIRETVFSYKNLIEGLE 24397
QY 209 ----- 208
Db 24398 YEFVRKCNLGESEWSEISEPITPKSDVPTOAPHKBEELNMLNVRYQSNATLVCKVGH 24457
QY 209 ----- 208
Db 24458 PKPIVKNYRQKEIADGLKYRIQEFKGGYHOLIIASVTDDDATVYQVRATNOGGSVST 24517
QY 209 ----- 208
Db 24518 ASLEVEVPAKIHLPKLTLEGMGAVHALRGEVSIKIPFSGKPOPVITWQKQDLIDNNGHY 24577
QY 209 ----- 208
Db 24578 QVIVTRSFVSLVFPNGVERKDGAFYVYCAKNRFGIDQKTVELDVADVPDPGRGVKVDVS 24637
QY 209 ----- 208
Db 24638 RDSVNLWTEPASDGGSKITNYIVEKCATTAERWLRVGOARETRYVINLFGKTSYQVRV 24697
QY 209 ----- 208
Db 24698 IAENKFGSKPSEPSEPTITTKEDKTRAMNYDEEVDRETVSMTKASHSSTKELYEKYMA 24757
QY 209 ----- 212
Db 24758 EDLGRGEFVHRCVETSSKTYMAKFKVKGTDQVLVKKEISILNIAHRNHLHESF 24817
QY 213 ----- 216
Db 24818 ESMEELVMIFEISOLDIFERINTSAFELNEREIVSVHVQVCEALQFLHSHNIGHDIRP 24877
QY 217 ----- 216
Db 24878 ENIIYOTRRSSTIKIEFGQARQLKPGDNFRLLFTAPEYAFVHQHVDVSTATDMSLG 24937
QY 217 ----- 218
Db 24938 TLVYVLLSGINPFLAETNQIIENIMNAEYTFDEEAFKEISIEAMDFVDRLLVKKRGRM 24997
QY 219 ----- 221
Db 24998 TASEALQHPWLKQKIERVSTKVIRTLKHRRYYHTLIKKDLNVMVSAARISCGGAIRSKG 25057
QY 222 LSL- ----- 225
Db 25058 VSAKVKVASIEIGPVSGQIMHAVGEGGHVKYCKIENYDQSTQVTVFGVRLNSEK 25117
QY 226 ----- 225
Db 25118 YEITYEDGVAILYKDKITKLDGTYRCKVNDYGEDSSYAEFLVKGVREVYDYCRRTMK 25177
QY 226 ----- 225
Db 25178 KIKRRTDMRLLEPPEFTPLYNKTA VGENVREGVTVTHPEPHVTWYKSGOKIKPGD 25237
QY 226 ----- 225
Db 25238 NDKYTFESDKGLYQLTJNSVTDDDAEYTVARNKYGEDSKAKLTVTLHPPPTDSTLR 25297
QY 226 ----- 225
Db 25298 PMFKRLLANABEQEQSVCFEIRVSGIPPTLKWEKQPLSLGNIEIIEHGLDYALH 25357

GenCore version 5.1.4_p5_4578
Copyright (c) 1993-2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:21 ; Search time 31.0588 seconds
(without alignments)
1132.630 Million cell updates/sec

Title: 2LINK7LINK7

Perfect score: 1430

Sequence: 1 MDKTHTCPCPAPPELLGSPS.....AFDWLKAIFYDKVAEKLKEAF 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1260	88.1	729	22 AAM52158	Humanised HMFG-1 h
2	1260	88.1	739	22 AAM52161	Humanised HMFG-1 h
3	1257	87.9	731	22 AAM52156	Humanised HMFG-1 h
4	1257	87.9	741	22 AAM52159	Humanised HMFG-1 h
5	1253	87.6	689	21 AAY96780	Ang-2-FD-Fc-FD fus
6	1251	87.5	652	19 AAW48650	Heavy chain of hma
7	1250	87.4	423	21 AAE28693	Fc-huAGP-1 (114-28
8	1250	87.4	441	21 AAE28692	Fc-huAGP-1 (95-281
9	1250	87.4	730	22 AAM52157	Humanised HMFG-1 h
10	1250	87.4	740	22 AAM52160	Humanised HMFG-1 h

11	1249	87.3	633	21 AAY84965	Amino acid sequenc
12	1247	87.2	426	21 AAB28695	Fc-huAGP-1 (120-29
13	1247	87.2	448	21 AAB28694	Fc-huAGP-1 (99-291
14	1246	87.1	247	21 AAB16958	Fc-TMP protein seq
15	1246	87.1	247	23 AAB73411	Fc-TPO mimetic pep
16	1246	87.1	248	21 AAB71953	Fc-IL-1 antagonist
17	1246	87.1	248	23 AAB73421	Fc-Interleukin 1 (
18	1246	87.1	268	21 AAB16959	Fc-TMP-TMP protein
19	1246	87.1	268	23 AAB73412	Fc-TMP-TMP amino a
20	1246	87.1	269	21 AAY96531	Human IgG1 Fc TMP
21	1246	87.1	374	19 AAW83963	Recombinant human
22	1246	87.1	374	19 AAW49075	Recombinant human
23	1246	87.1	401	22 AAY72922	Human met-Fc (lack
24	1246	87.1	401	22 AAB80904	Human metFcdeltaC-
25	1246	87.1	689	21 AAY96779	Ang-1-FD-Fc-FD fus
26	1245	87.1	252	21 AAB17955	Fc-VEGF antagonist
27	1245	87.1	252	23 AAB73423	Fc-VEGF antagonist
28	1245	87.1	651	18 AAW26649	Chimeric receptor
29	1245	87.1	692	18 AAW26650	Chimeric receptor
30	1241	86.8	379	19 AAW83962	Recombinant human
31	1241	86.8	379	19 AAW49073	Recombinant human
32	1241	86.8	379	19 AAW49074	Human IgG1 Fc-inte
33	1240	86.7	389	22 AAW50247	Ganglioside GD2 sp
34	1240	86.7	581	22 AAB81972	Ganglioside GD3 sp
35	1240	86.7	582	22 AAB81987	Ganglioside GD3 sp
36	1240	86.7	582	22 AAB81991	Ganglioside GM2 an
37	1240	86.7	583	22 AAB83156	Fc-TNF-alpha inhib
38	1239	86.6	248	21 AAB17951	Fc-TNF-alpha inhib
39	1239	86.6	248	23 AAB73419	Fc-EPO protein seq
40	1239	86.6	253	21 AAB16964	Fc-EPO mimetic pep
41	1239	86.6	253	23 AAB73415	Fc-EPO protein
42	1239	86.6	277	21 AAB16967	Fc-EPO protein
43	1239	86.6	277	23 AAB73418	Fc-EPO protein
44	1239	86.6	660	20 AAX13463	Ectromelia A39R se
45	1239	86.6	660	21 AAB28523	Ectromelia A39R se

ALIGNMENTS

RESULT 1

AAM52158

ID AAM52158 standard; Protein; 729 AA.

AC AAM52158;

XX AAM52158;

DT 05-FEB-2002 (first entry)

XX Humanised HMFG-1 heavy chain/DNase I fusion protein 3.

DE DE

XX Humanised monoclonal antibody; polymorphic epithelial mucin; PEM1;

KW cytotoxic; endonuclease; DNase I; human; cytostatic; cancer; apoptosis.

XX Homo sapiens.

OS Synthetic.

XX WO200174905-A1.

PN 11-OCT-2001.

XX 26-MAR-2001; 2001WO-GB01324.

XX 03-APR-2000; 2000GB-0008049.

PR 02-OCT-2000; 2000US-237159P.

XX (ANTI-) ANTISOMA RES LTD.

XX Young RJ;

XX WPI; 2001-662969/76.

XX Novel compound used to treat cancer has target cell-specific portion

XX comprising humanised monoclonal antibody having specificity for

PT

```
PT polymorphic epithelial mucin, and cytotoxic portion having
XX endonucleolytic activity
PS Claim 20; Figure 9; 176pp; English.
XX
CC The invention relates to a compound which comprises a target
CC cell-specific portion, comprising an humanised monoclonal antibody,
CC having specificity for polymorphic epithelial mucin (PEM) or its antigen
CC binding fragment and a cytotoxic portion having endonucleolytic activity,
CC exemplified by AAM52154-AAM52168 and encoded by ABA02682-ABA02728. The
CC compound has cytostatic activity useful for treating cancer and acting as
CC a potential inducer of apoptosis.
XX
SQ Sequence 729 AA;
Query Match 88.1%; Score 1260; DB 22; Length 729;
Best Local Similarity 50.5%; Pred. No. 3.7e-33;
Matches 241; Conservative 6; Mismatches 15; Indels 215; Gaps 5;
QY 2 DKHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61
Db 241 DKHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 300
QY 62 GVEVHNKTKPREEQNSTYRVSVLTCLVKGFPSPDIAVEWESNGOPENNYKTTTPVLD 121
Db 301 GVEVHNKTKPREEQNSTYRVSVLTCLVKGFPSPDIAVEWESNGOPENNYKTTTPVLD 360
QY 122 GPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLD 181
Db 361 GPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLD 420
QY 182 DGSFFLYSKLTVDKSRWQOGNFSVMSVMEALHNHYTQKSLSLSPGKWLK----- 232
Db 421 DGSFFLYSKLTVDKSRWQOGNFSVMSVMEALHNHYTQKSLSLSPGKWLK----- 480
QY 233 -----KAFYDKV-----AEKLKEA 263
Db 661 FQWLIPDSADTTATPTHCAIDRIYVAGMLLRGAVVPDSALPFNFQAAYGLSDQLAQ 717
RESULT 2
AAM52161
ID AAM52161 standard; Protein; 739 AA.
XX
AC AAM52161;
XX
DT 05-FEB-2002 (first entry)
XX
DE Humanised HMFG-1 heavy chain/DNase I fusion protein 6.
XX
KW Humanised monoclonal antibody; polymorphic epithelial mucin; PEM;
KW cytotoxic; endonuclease; DNase I; human; cytostatic; cancer; apoptosis.
XX
OS Homo sapiens.
XX
OS Synthetic.
XX
PN WO200174905-A1.
XX
PD 11-OCT-2001.
XX
PF 26-MAR-2001; 2001WO-GB01324.
```

```
XX
PR 03-APR-2000; 2000GB-0008049.
PR 02-OCT-2000; 2000US-237159P.
XX
PA (ANTI-) ANTISOMA RES LTD.
XX
PI Young RJ;
XX
DR WPI; 2001-662969/76.
XX
PT Novel compound used to treat cancer has target cell-specific portion
PT comprising humanised monoclonal antibody having specificity for
PT polymorphic epithelial mucin, and cytotoxic portion having
PT endonucleolytic activity
XX
PS Claim 20; Figure 12; 176pp; English.
XX
CC The invention relates to a compound which comprises a target
CC cell-specific portion, comprising an humanised monoclonal antibody,
CC having specificity for polymorphic epithelial mucin (PEM) or its antigen
CC binding fragment and a cytotoxic portion having endonucleolytic activity,
CC exemplified by AAM52154-AAM52168 and encoded by ABA02682-ABA02728. The
CC compound has cytostatic activity useful for treating cancer and acting as
CC a potential inducer of apoptosis.
XX
SQ Sequence 739 AA;
Query Match 88.1%; Score 1260; DB 22; Length 739;
Best Local Similarity 50.5%; Pred. No. 3.7e-33;
Matches 241; Conservative 6; Mismatches 15; Indels 215; Gaps 5;
QY 2 DKHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61
Db 241 DKHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 300
QY 62 GVEVHNKTKPREEQNSTYRVSVLTCLVKGFPSPDIAVEWESNGOPENNYKTTTPVLD 121
Db 301 GVEVHNKTKPREEQNSTYRVSVLTCLVKGFPSPDIAVEWESNGOPENNYKTTTPVLD 360
QY 122 GPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLD 181
Db 361 GPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGOPENNYKTTTPVLD 420
QY 182 DGSFFLYSKLTVDKSRWQOGNFSVMSVMEALHNHYTQKSLSLSPGKWLK----- 232
Db 421 DGSFFLYSKLTVDKSRWQOGNFSVMSVMEALHNHYTQKSLSLSPGKWLK----- 480
QY 233 -----KAFYDKV-----AEKLKEA 263
Db 481 GETKMSNATLVSYIVQILSRDIALVQEVDRSHLTAVGKLLDNLDNODAPDTHYVYVSEPL 540
QY 233 ----- 232
Db 541 GRNSYKERYLFVYRPDQVSAVDYDDGCEPCGNDTFNREPAIVRFFSRFTVREFAIV 600
QY 233 -----AFYD---KVAEK-----LKEA 245
Db 601 PLHAAPGDAVEIDALDYLDVQEKWGLEVDVLMGDFNAGGSYVRPQWSSIRLWTSPT 660
QY 246 FQWL-----KAFYDKV-----AEKLKEA 263
Db 661 FQWLIPDSADTTATPTHCAIDRIYVAGMLLRGAVVPDSALPFNFQAAYGLSDQLAQ 717
RESULT 3
AAM52156
ID AAM52156 standard; Protein; 731 AA.
XX
AC AAM52156;
XX
DT 05-FEB-2002 (first entry)
XX
DE Humanised HMFG-1 heavy chain/DNase I fusion protein 1.
```

XX Humanised monoclonal antibody; polymorphic epithelial mucin; PEM1;
KW cytotoxic; endonuclease; DNase I; human; cytostatic; cancer; apoptosis.
XX Homo sapiens.
OS Synthetic.
XX WO200174905-A1.
XX 11-OCT-2001.
XX 26-MAR-2001; 2001WO-GB01324.
XX 03-APR-2000; 2000GB-0008049.
PR 02-OCT-2000; 2000US-237159P.
XX (ANTI-) ANTISOMA RES LTD.
XX Young RJ;
XX WPI; 2001-662969/76.
XX Novel compound used to treat cancer has target cell-specific portion
PT comprising humanised monoclonal antibody having specificity for
PT polymorphic epithelial mucin, and cytotoxic portion having
PT endonucleolytic activity
XX Claim 20; Figure 7; 176pp; English.
XX The invention relates to a compound which comprises a target
CC cell-specific portion, comprising an humanised monoclonal antibody,
CC having specificity for polymorphic epithelial mucin (PEM) or its antigen
CC binding fragment and a cytotoxic portion having endonucleolytic activity,
CC exemplified by AAM52154-AM52168 and encoded by ABA02682-ABA02728. The
CC compound has cytostatic activity useful for treating cancer and acting as
CC a potential inducer of apoptosis.
XX SQ Sequence 731 AA;
Query Match 87.9%; Score 1257; DB 22; Length 731;
Best Local Similarity 50.1%; Pred. No. 4.6e-33;
Matches 240; Conservative 6; Mismatches 16; Indels 217; Gaps 4;
QY 2 DKHTCPCPAPELLGGPSVFLPFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNMYVD 61
Db 241 DKHTCPCPAPELLGGPSVFLPFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNMYVD 300
QY 62 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
Db 301 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 360
QY 122 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 181
Db 361 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 420
QY 182 DGSFFLYSKLTVDKSRWQOGNVFSCSVMEALHNHYTQKSLSLSPGK----- 228
Db 421 DGSFFLYSKLTVDKSRWQOGNVFSCSVMEALHNHYTQKSLSLSPGK----- 480
QY 229 ----- 228
Db 481 TFGETKMSNATLSYIVQILSRDIALVQEVDRSHLTVAGKLLDNLNQADPTYYHYVSE 540
QY 229 ----- 228
Db 541 PLGRNSYKERYLFYRPDPQVSAVDSYDDGCEPCGNDTFNREPAIVRFRSRTVEFEA 600
QY 229 -----DWLKAFYDKVAEK-----LK 243
Db 601 IVPLHAAPGDAVEIDALYDLDVQEKWGLDVMGLMGDFNAGCSYVRPQWSQIRLWTS 660
QY 244 EAFDWL-----KAFYDKV-----AEKLKEA 263
| | | | |
| | | | |

Db 661 PTFOWLIPDSADTTATPTTHCAVDRIYVAGMLLRGAVVPDSALPFFQAAVGLSDQLAQA 719
RESULT 4
AAM52159
ID AAM52159 standard; Protein; 741 AA.
XX AC AAM52159;
XX DT 05-FEB-2002 (first entry)
XX DE Humanised HMFG-1 heavy chain/DNase I fusion protein 4.
XX KW Humanised monoclonal antibody; polymorphic epithelial mucin; PEM1;
XX cytotoxic; endonuclease; DNase I; human; cytostatic; cancer; apoptosis.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN WO200174905-A1.
XX PD 11-OCT-2001.
XX PF 26-MAR-2001; 2001WO-GB01324.
XX PR 03-APR-2000; 2000GB-0008049.
PR 02-OCT-2000; 2000US-237159P.
XX (ANTI-) ANTISOMA RES LTD.
XX Young RJ;
XX WPI; 2001-662969/76.
XX Novel compound used to treat cancer has target cell-specific portion
PT comprising humanised monoclonal antibody having specificity for
PT polymorphic epithelial mucin, and cytotoxic portion having
PT endonucleolytic activity
XX Claim 20; Figure 10; 176pp; English.
XX The invention relates to a compound which comprises a target
CC cell-specific portion, comprising an humanised monoclonal antibody,
CC having specificity for polymorphic epithelial mucin (PEM) or its antigen
CC binding fragment and a cytotoxic portion having endonucleolytic activity,
CC exemplified by AAM52154-AM52168 and encoded by ABA02682-ABA02728. The
CC compound has cytostatic activity useful for treating cancer and acting as
CC a potential inducer of apoptosis.
XX SQ Sequence 741 AA;
Query Match 87.9%; Score 1257; DB 22; Length 741;
Best Local Similarity 50.1%; Pred. No. 4.7e-33;
Matches 240; Conservative 6; Mismatches 16; Indels 217; Gaps 4;
QY 2 DKHTCPCPAPELLGGPSVFLPFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNMYVD 61
Db 241 DKHTCPCPAPELLGGPSVFLPFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNMYVD 300
QY 62 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
Db 301 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 360
QY 122 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 181
Db 361 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 420
QY 182 DGSFFLYSKLTVDKSRWQOGNVFSCSVMEALHNHYTQKSLSLSPGK----- 228
Db 421 DGSFFLYSKLTVDKSRWQOGNVFSCSVMEALHNHYTQKSLSLSPGK----- 480
QY 229 ----- 228

Db 481 TFGETKMSNATLYSYIVQILSRDYDIALVOEVRDLSHTAVGKLLDNLNQADPDTHYHVWSE 540
 QY 229 ----- 228
 Db 541 PLGRNSYKERYLFVYRPDQVSAVDSYVDDGCEPCGNDTFNREPAIVRFESRETEVREFA 600
 QY 229 -----DWLKAFYDKVAEK-----LJK 243
 Db 601 IVPLHAAPGDAVAEIDALYDVLDOVKWGLDVMGLMGDFNAGCSYVRPSSIRLWTS 660
 QY 244 EAFDWL-----KAFYDKV-----AEKLEKA 263
 Db 661 PTQWLIPDSADTATPTTCANDRIVVAGMLLRGAVVPDSALPFFNQAAAYGLSDQLAQA 719
 RESULT 5
 AAY96780
 ID AAY96780 standard; Protein; 689 AA.
 AC AAY96780;
 DT 26-SEP-2000 (first entry)
 XX Ang-2-FD-Fc fusion protein.
 DE
 DE Angiopoietin-2; Tie-2 receptor; ligand; antagonist; fibrinogen domain;
 KW dimer; Fc domain; fusion protein; anti-proliferative; cytostatic;
 KW leukaemia; haematopoietic factor.
 XX
 OS Chimeric - Homo sapiens.
 OS Chimeric - Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..15
 FT /label= Trypsin_signal_sequence
 FT Domain 16..230
 FT /label= Ang-2_fibrinogen-like_domain
 FT Peptide 231..234
 FT /label= Bridging_peptide
 FT Region 235..466
 FT /label= Human_IgG1_Fc_region
 FT Peptide 467..474
 FT /label= Bridging_peptide
 FT Domain 475..689
 FT /label= Ang-2_fibrinogen-like_domain
 XX
 PN WO200037642-A1.
 XX
 PD 29-JUN-2000.
 XX
 PF 23-DEC-1999; 99WO-US30900.
 XX
 PR 23-DEC-1998; 98US-0113387.
 XX
 PA (REG-) REGENERON PHARM INC.
 XX
 PI Davis SJ, Gale NW, Yancopoulos GD, Stahl N;
 XX WPI; 2000-442670/38.
 DR N-PSDB; AAA51344.
 XX
 PT Polynucleotide encoding a fusion polypeptide, useful for promoting
 PT differential function and influencing phenotype, comprises two subunits
 PT containing at least one copy of the receptor binding domain of a ligand
 XX
 PS Example 1; Fig 4A-E; 97pp; English.
 XX
 CC Angiopoietin-1 (Ang-1) is a Tie-2 receptor ligand and is an agonist for
 CC Tie-2, whereas angiopoietin-2 (Ang-2) is a naturally occurring antagonist
 CC of the Tie-2 receptor. The fibrinogen domains (FD) of Ang-1 and Ang-2 are
 CC the receptor-binding domains and dimerized versions, of e.g. Ang-1-FD-Fc
 CC (Ang-1 fibrinogen domain fused to an Fc domain), can bind to the Tie-2
 CC receptor with much higher affinity than monomeric Ang-1-FD (dimerization

CC occurs between the Fc components of adjacent molecules). However,
 CC Ang-1-FD-Fc is not able to induce phosphorylation (activate) the Tie-2
 CC receptor on endothelial cells unless it is further clustered with goat
 CC anti-human Fc antibodies. The novel fusion proteins, mutant versions of
 CC Ang-1-FD and Ang-2-FD, were designed that were intrinsically more highly
 CC clustered. Tie-2 agonist fusion proteins may be used as haematopoietic
 CC factors. Tie-2 receptor antagonist fusion proteins may be used to
 CC diagnose or treat, e.g. myeloproliferative or other proliferative
 CC disorders of blood forming organs, e.g. thrombocythemia, polycythemia
 CC and leukemias. Production of homogeneous forms of clustered ligands is
 CC broadly applicable to improve the affinity and/or increase the activity
 CC of a ligand as compared to the native form of the ligand. Ephrin fusion
 CC proteins have also been constructed, which may be useful for treating
 CC neurological disorders. The ephrin fusion proteins are preferably capable
 CC of binding to Elk receptor and are especially Efl-6 antagonists.
 XX
 SQ Sequence 689 AA;

Query Match 87.6%; Score 1253; DB 21; Length 689;
 Best Local Similarity 55.1%; Pred. No. 5.4e-33;
 Matches 238; Conservative 7; Mismatches 9; Indels 178; Gaps 4;

QY 2 DKTHCPCPAPPELLGSPVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKENWYVD 61
 Db 240 DKTHCPCPAPPELLGSPVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKENWYVD 299
 QY 62 GVEVHNATKPREEQNSTYRVVSVLTVLHODWLNKGYCKVSNKALPAPIETISKAK 121
 Db 300 GVEVHNATKPREEQNSTYRVVSVLTVLHODWLNKGYCKVSNKALPAPIETISKAK 359
 QY 122 GQPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 181
 Db 360 GQPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 419
 QY 182 DGSFELYSLKLVDSRWQGNVFCVSMHEALHNNHYTKSLSPGK----- 228
 Db 420 DGSFELYSLKLVDSRWQGNVFCVSMHEALHNNHYTKSLSPGKGGSGAPRDCAE 479
 QY 229 -----DWLKAFYDKVA-----EKLKEAF 246
 Db 480 VFKSGHTTNGIYTLTFPNSTEIKAYCDMEAGGGWTIIQRDGSVDVDFQRTWEYKVG 539
 QY 247 -----DW----- 248
 Db 540 GNPGEYWLGNFVSQLTNQORYVLKIHLDWEGNEAYSLYEHFVLSSELNRYRIHLKGL 599
 QY 249 ----- 248
 Db 600 TGTAGKISSISOPGNDFSTKGDNDKICKCSQMLTGGWFDACGSPSNLNGMYYPQRONT 659
 QY 249 -----LKAFYDK 255
 Db 660 NKENGIRKYYWK 671

RESULT 6
 AAW48650
 ID AAW48650 standard; Protein; 652 AA.
 XX
 AC AAW48650;
 XX
 DT 04-AUG-1998 (first entry)
 XX Heavy chain of hmAb425 fused to TNF alpha.
 DE
 DE Antibody-cytokine fusion protein; tricistronic vector; chimeric;
 KW TNF alpha; IL-2; IRES; internal ribosome entry site.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Region 1..494

FT /note= "Heavy chain of human mAb 425"
 FT 495...652
 FT /note= "TNF alpha"
 XX
 PN W09811241-A1.
 XX
 PD 19-MAR-1998.
 XX
 PF 02-SEP-1997; 97WO-EP04765.
 XX
 PR 30-SEP-1996; 96EP-0115635.
 PR 16-SEP-1996; 96EP-0114820.
 XX
 XX (MERE) MERCK PATENT GMBH.
 PA
 XX Bruenner W, Burge C, Dunker R, Hauser H, Mielke C;
 PI Rieke E, Von Hoegen I, Welge T;
 PI
 XX WPI; 1998-207400/18.
 DR N-PSDB; AAV18096.
 DR
 XX
 XX Oligo:cistronic expression vector - useful for production of, e.g.
 PT MAb425/TNF-a or MAb425/IL-2 antibody fusion protein
 PT
 XX Disclosure; Fig 15; 89pp; English.
 PS
 XX
 CC The present sequence represents a fusion protein comprising of TNF
 CC alpha fused to the C-terminus of the heavy chain of the human
 CC monoclonal antibody 425 (hmb425). The hmb425 has specificity for
 CC the human EGF receptor. The invention claims for a new pmCLDHAP
 CC tricistronic vector (AAV18096) for the expression of an
 CC antibody-cytokine fusion protein, hmb425-TNF alpha. The TNF alpha
 CC sequence can be substituted by the IL-2 sequence. The vector also
 CC contains a strong promoter/enhancer unit, a selection marker gene and at
 CC least two poliovirus derived internal ribosomal entry site (IRES)
 CC sequences. The vector can be expressed in mammalian host cells for the
 CC production of heteromeric fusion proteins. This expression system is
 CC claimed to produce the heteromeric proteins in high yields.
 XX
 SQ Sequence 652 AA;

Query Match 87.5%; Score 1251; DB 19; Length 652;
 Best Local Similarity 65.4%; Pred. No. 5.6e-33;
 Matches 238; Conservative 8; Mismatches 13; Indels 105; Gaps 5;

QY 2 DKHTCCPCPAPELLGGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVD 61
 DB 268 DKHTCCPCPAPELLGGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVD 327
 QY 62 GVEVHNKTRPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIETISKAK 121
 DB 328 GVEVHNKTRPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIETISKAK 387
 QY 122 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 181
 DB 388 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 447
 QY 182 DGSFFLYSKLTVDKSRWQOGNVFSCSVHMEALHNHYTQKSLSLSPGK----- 228
 DB 448 DGSFFLYSKLTVDKSRWQOGNVFSCSVHMEALHNHYTQKSLSLSPGKVRSSRTSPDKP 507
 QY 229 -----DWLK-----AFYDKVAEK----- 241
 DB 508 VAHVANPQAEQQLWLNRRNALLANGVELNQLVVPSEGLYLYISQVLFKGGCPST 567
 QY 242 -----LK-----EAFDWLKAFY-----DKV 256
 DB 568 HVLTLTISIAVSQYTKVNLSSAISKPCQRETPEGAEPWYEPYILGGVFWFQLEKGDRL 627
 QY 257 AEKL 260
 DB 628 SAEI 631

RESULT 7
 AAB28693
 ID AAB28693 standard; Protein; 423 AA.
 XX
 AC AAB28693;
 XX
 DT 14-FEB-2001 (first entry)
 XX
 DE Fc-huAGP-1 (114-281) fusion protein.
 XX
 KW Human; AGP-1; type II transmembrane protein; cytostatic; antiviral;
 KW antiinflammatory; hepatotropic; antiarteriosclerotic; anti-HIV; HIV;
 KW human immunodeficiency virus; apoptosis; proliferative disorder;
 KW cancer; hepatitis; acquired immunodeficiency syndrome; AIDS;
 KW autoimmune disorder; transplant rejection; cardiovascular disease;
 KW arteriosclerosis; Fc-huAGP-1; fusion protein.
 XX
 OS Homo sapiens.
 XX
 PN W0200063253-A1.
 XX
 PD 26-OCT-2000.
 XX
 PF 24-MAR-2000; 2000WO-US08004.
 XX
 PR 16-APR-1999; 99US-0293245.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Hsu H, Meng S;
 XX
 XX WPI; 2000-665240/64.
 DR N-PSDB; AAC67833.
 DR
 XX Fusion protein of AGP-1 protein and an Fc region, used to treat
 PT proliferative disorders, immune disorders, and virally-induced
 PT disorders -
 PT
 XX Disclosure; Fig 4; 93pp; English.
 XX
 CC The present sequence is an AGP-1 fusion protein. AGP-1 is a
 CC type II transmembrane protein. The fusion proteins comprise an Fc
 CC immunoglobulin region fused to the N-terminal portion of the AGP-1
 CC protein. The fusion proteins can be used to induce apoptosis in a tissue,
 CC and to treat proliferative disorders, immune disorders, or
 CC virally-induced disorders. The proliferative disorders include cancers,
 CC such as breast, prostate, lung or colon cancer. The viral infections
 CC include hepatitis, and acquired immunodeficiency syndrome (AIDS), and the
 CC immune disorders may be autoimmune disorders or transplant rejection.
 CC Cardiovascular diseases such as arteriosclerosis may also be treated. The
 CC AGP-1 containing fusion proteins have increased biological activity
 CC compared to the soluble AGP-1 proteins used in prior art therapies.
 XX
 SQ Sequence 423 AA;

Query Match 87.4%; Score 1250; DB 21; Length 423;
 Best Local Similarity 65.6%; Pred. No. 2.2e-33;
 Matches 240; Conservative 6; Mismatches 15; Indels 105; Gaps 5;

QY 2 DKHTCCPCPAPELLGGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVD 61
 DB 29 DKHTCCPCPAPELLGGPSVFLPFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVD 88
 QY 62 GVEVHNKTRPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIETISKAK 121
 DB 89 GVEVHNKTRPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIETISKAK 148
 QY 122 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 181
 DB 149 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 208
 QY 182 DGSFFLYSKLTVDKSRWQOGNVFSCSVHMEALHNHYTQKSLSLSPGK----- 228

```
Db 209 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGKVRGPRVAHHI 268
QY 229 -----DW-----LKAFFY----- 235
Db 269 TGTGRSNTLSSPNSKNEKALGRKINSWESSRSGHSLNLHURNGELVIHEKGFYIYS 328
QY 236 -----DKVAEKLKEAFDWLKAFY-----DKVAE----- 258
Db 329 QTVFRFOEIKENTKNDKQWQVYIYKVTSPDPILLMKMSARNSCWSKDAEYGLYSIQGG 388
QY 259 --KLKE 262
Db 389 IFELKE 394

RESULT 8
AAB28692
ID AAB28692 standard; Protein; 441 AA.
XX
AC AAB28692;
XX
DT 14-FEB-2001 (first entry)
XX
DE Fc-huAGP-1 (95-281) fusion protein.
XX
KW Human; AGP-1; type II transmembrane protein; cytostatic; antiviral;
KW antiinflammatory; hepatotropic; antiarteriosclerotic; anti-HIV; HIV;
KW human immunodeficiency virus; apoptosis; proliferative disorder;
KW cancer; hepatitis; acquired immunodeficiency syndrome; AIDS;
KW autoimmune disorder; transplant rejection; cardiovascular disease;
KW arteriosclerosis; Fc-huAGP-1; fusion protein.
XX
OS Homo sapiens.
XX
PN WO200063253-A1.
XX
PD 26-OCT-2000.
XX
PF 24-MAR-2000; 2000WO-US08004.
XX
PR 16-APR-1999; 99US-0293245.
XX
PA (AMGE-) AMGEN INC.
XX
PI Hsu H, Meng S;
XX
DR WPI; 2000-665240/64.
XX
DR N-PSDB; AAC67832.
XX
PT Fusion protein of AGP-1 protein and an Fc region, used to treat
PT proliferative disorders, immune disorders, and virally-induced
PT disorders -
XX
PS Disclosure; Fig 3; 93pp; English.
XX
CC The present sequence is an AGP-1 fusion protein. AGP-1 is a
CC type II transmembrane protein. The fusion proteins comprise an Fc
CC immunoglobulin region fused to the N-terminal portion of the AGP-1
CC protein. The fusion proteins can be used to induce apoptosis in a tissue,
CC and to treat proliferative disorders, immune disorders, or
CC virally-induced disorders. The proliferative disorders include cancers,
CC such as breast, prostate, lung or colon cancer. The viral infections
CC include hepatitis, and acquired immunodeficiency syndrome (AIDS), and the
CC immune disorders may be autoimmune disorders or transplant rejection.
CC Cardiovascular diseases such as arteriosclerosis may also be treated. The
CC AGP-1 containing fusion proteins have increased biological activity
CC compared to the soluble AGP-1 proteins used in prior art therapies.
XX
SQ Sequence 441 AA;
Query Match 87.4%; Score 1250; DB 21; Length 441;
Best Local Similarity 62.3%; Pred. No. 2.4e-33;
```

```
Matches 240; Conservative 6; Mismatches 15; Indels 124; Gaps 5;
QY 2 DKTHTCPPCPAPPELLGGPSVFLFPPPKPDTLMTSRTPETCVVVDVSHEDPEVKFNWYD 61
Db 29 DKTHTCPPCPAPPELLGGPSVFLFPPPKPDTLMTSRTPETCVVVDVSHEDPEVKFNWYD 88
QY 62 GVEVHNATKPREEQYNSTYRVSVLTVLHQDLNGLNGKEYCKVSNKALPAPIEKTISKAK 121
Db 89 GVEVHNATKPREEQYNSTYRVSVLTVLHQDLNGLNGKEYCKVSNKALPAPIEKTISKAK 148
QY 122 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGYGSPDSIAVWESNGOPENNYKTTTPVLD 181
Db 149 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGYGSPDSIAVWESNGOPENNYKTTTPVLD 208
QY 182 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGK----- 228
Db 209 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGKSEETISTVQEQK 268
QY 229 -----DW----- 230
Db 269 QNISPLVREGRGPRVAHHITGRSNTLSSPNSKNEKALGRKINSWESSRSGHSLNL 328
QY 231 -----LKAFFY-----DKVAEKLKEAFDWLKAFY----- 253
Db 329 HLRNGELVIHEKGFYIYSQTVFRFOEIKENTKNDKQWQVYIYKVTSPDPILLMKMSAR 388
QY 254 ----DKVAE-----KLKE 262
Db 389 NSCWSKDAEYGLYSIQGGIFELKE 413

RESULT 9
AAM52157
ID AAM52157 standard; Protein; 730 AA.
XX
AC AAM52157;
XX
DT 05-FEB-2002 (first entry)
XX
DE Humanised HMEF-1 heavy chain/DNase I fusion protein 2.
XX
KW Humanised monoclonal antibody; polymorphic epithelial mucin; PEM1;
KW cytotoxic; endonuclease; DNase I; human; cytostatic; cancer; apoptosis.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200174905-A1.
XX
PD 11-OCT-2001.
XX
PF 26-MAR-2001; 2001WO-GB01324.
XX
PR 03-APR-2000; 2000GB-0008049.
XX
PR 02-OCT-2000; 2000US-237159P.
XX
PA (ANTI-) ANTISOMA RES LTD.
XX
PI Young RJ;
XX
WPI; 2001-662969/76.
XX
Novel compound used to treat cancer has target cell-specific portion
comprising humanised monoclonal antibody having specificity for
polymorphic epithelial mucin, and cytotoxic portion having
endonucleolytic activity -
XX
PS Claim 20; Figure 8; 176pp; English.
XX
CC The invention relates to a compound which comprises a target
CC cell-specific portion, comprising an humanised monoclonal antibody,
CC having specificity for polymorphic epithelial mucin (PEM) or its antigen
CC binding fragment and a cytotoxic portion having endonucleolytic activity,
```


FT	Peptide	27..86
FT	/note= "signal peptide from murine T86.66 antibody	
FT	kappa light chain"	
FT	Region	21..126
FT	/note= "kappa light chain"	
FT	Peptide	145..266
FT	/note= "anti-CD20 variable regions"	
FT	Region	283..392
FT	/note= "GS18 linker"	
FT	Region	393..499
FT	/note= "hinge region"	
FT	Region	500..521
FT	/note= "CH3 region"	
FT	Region	522..633
FT	/note= "CD4 transmembrane region"	
XX	/note= "zeta chain"	
XX	WO200023573-A2.	
XX	27-APR-2000.	
XX	PD	
XX	PF	
XX	20-OCT-1999;	99WO-US24484.
XX	PR	
XX	20-OCT-1998;	98US-0105014.
XX	(CITY) CITY OF HOPE.	
PA	Raubitschek A, Jensen MC, Wu AM;	
PI		
DR	WPI; 2000-339676/29.	
XX	N-PSDB; AAA15019.	
XX	Genetically engineered CD20-specific redirected T cells useful for	
PT	treating a CD20+ malignancy, such as non-Hodgkin's lymphoma or CD20+	
PT	acute or chronic leukemia, and autoimmune disease -	
XX	Claim 10; Page 53-55; 58pp; English.	
CC	The present sequence represents a synthetic CD20-specific chimeric	
CC	receptor. The specification describes CD-20 specific redirected T cells	
CC	which express and bear on the cell surface membrane a CD20-chimeric	
CC	receptor comprising an intracellular signalling domain, a transmembrane	
CC	domain and an extracellular domain, the extracellular domain comprising	
CC	a CD20-specific receptor. The genetically engineered CD20-specific	
CC	redirected T cells are useful for treating a CD20+ malignancy, such	
CC	as non-Hodgkin's lymphoma or CD20+ acute or chronic leukemia, in a	
CC	human patient having previously undergone myeloablative chemotherapy and	
CC	stem cell rescue. The genetically engineered CD20-specific redirected	
CC	T cells are also useful for abrogating an untoward B cell function, such	
CC	as autoimmune disease (lupus or rheumatoid arthritis) in a patient.	
XX		
SQ	Sequence 633 AA;	
	Query Match 87.3%; Score 1249; DB 21; Length 633;	
	Best Local Similarity 73.9%; Pred. No. 6.1e-33;	
	Matches 238; Conservative 7; Mismatches 16; Indels 61; Gaps	
QY	2 DKHTCPPCPAPELLGSPVFLPPPKDPTLMISRTPEVTVCVVVDVSHEDPEVKFNWYD 61	
Db		
	273 DKHTCPPCPAPELLGSPVFLPPPKDPTLMISRTPEVTVCVVVDVSHEDPEVKFNWYD 332	
QY	62 GVEVHNKTKPREQQNSTYRVSVLTVLHQDWLNCKEYCKCVSNKALPAPIETISKAK 121	
Db		
	333 GVEVHNKTKPREQQNSTYRVSVLTVLHQDWLNCKEYCKCVSNKALPAPIETISKAK 392	
QY	122 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSPDIAEVESNGOPENNYKTTPPVLD 181	
Db		
	393 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSPDIAEVESNGOPENNYKTTPPVLD 452	
QY	182 DGSFELYSKLTVDKSRWQGNQFSCVMHEALNHHYTQKLSLSFGCKDWL----- 231	
Db		
	453 DGSFFLYSKLTVDKSRWQGNQFSCVMHEALNHHYTQKLSLSFGCKMWGLLVLGAGLL .512	
QY	232 -----KATYDKVAEKLEKFAFDWL----- 249	

Db 89 GVEVHNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 148
 QY 122 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 181
 Db 149 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 208
 QY 182 DGSFFLYSKLTVDKSRWQGNVFSCSYVMEALHNHYTQKSLSLSPGK----- 228
 Db 209 DGSFFLYSKLTVDKSRWQGNVFSCSYVMEALHNHYTQKSLSLSPGKVAAHITG 268
 QY 229 -----DW-----LKAEDYDKV 238
 Db 269 ITRRSNALIPISKDGLTLCQKIESWESSRKGHSFLNHVLFNRGELVIEQGLYYISQT 328
 QY 239 AEKLKEAFDWLK-AFYDKVAEK-----LKEA 263
 Db 329 YFRFOEAEDASKMYSKDKVTKQLVQYIYKTSYPDPPIVLMKSA 372

RESULT 13
 AAB28694
 ID AAB28694 standard; Protein; 448 AA.
 XX
 AC AAB28694;
 XX
 DT 14-FEB-2001 (first entry)
 DE FC-muAGP-1 (99-291) fusion protein.
 XX
 KW Mouse; AGP-1; type II transmembrane protein; cytostatic; antiviral;
 KW antiinflammatory; hepatotropic; arteriosclerotic; anti-HIV; HIV;
 KW human immunodeficiency virus; apoptosis; proliferative disorder;
 KW cancer; hepatitis; acquired immunodeficiency syndrome; AIDS;
 KW autoimmune disorder; transplant rejection; cardiovascular disease;
 KW arteriosclerosis; FC-muAGP-1; fusion protein.
 XX
 OS Mus sp.
 XX
 PN WO200063253-A1.
 XX
 PD 26-OCT-2000.
 XX
 PF 24-MAR-2000; 2000WO-US08004.
 XX
 PR 16-APR-1999; 99US-0293245.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Hsu H, Meng S;
 XX
 DR WPI; 2000-665240/64.
 DR N-PSDB; AAC67834.
 XX
 PT Fusion protein of AGP-1 protein and an Fc region, used to treat
 PT proliferative disorders, immune disorders, and virally-induced
 PT disorders -
 XX
 PS Disclosure; Fig 5; 93pp; English.
 XX
 CC The present sequence is an AGP-1 fusion protein. AGP-1 is a
 CC type II transmembrane protein. The fusion proteins comprise an Fc
 CC immunoglobulin region fused to the N-terminal portion of the AGP-1
 CC protein. The fusion proteins can be used to induce apoptosis in a tissue,
 CC and to treat proliferative disorders, immune disorders, or
 CC virally-induced disorders. The proliferative disorders include cancers,
 CC such as breast, prostate, lung or colon cancer. The viral infections
 CC include hepatitis, and acquired immunodeficiency syndrome (AIDS), and the
 CC immune disorders may be autoimmune disorders or transplant rejection.
 CC Cardiovascular diseases such as arteriosclerosis may also be treated. The
 CC AGP-1 containing fusion proteins have increased biological activity
 CC compared to the soluble AGP-1 proteins used in prior art therapies.
 XX
 SQ Sequence 448 AA;

Query Match 87.2%; Score 1247; DB 21; Length 448;
 Best Local Similarity 65.8%; Pred. No. 3.2e-33;
 Matches 240; Conservative 4; Mismatches 18; Indels 103; Gaps 4;
 QY 2 DKTHTCPCPAPELLGSPVFLPPPKDLMISRTPEVTCVVVDVSHEDPEVKFNKYVD 61
 Db 29 DKTHTCPCPAPELLGSPVFLPPPKDLMISRTPEVTCVVVDVSHEDPEVKFNKYVD 88
 QY 62 GVEVHNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
 Db 89 GVEVHNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 148
 QY 122 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 181
 Db 149 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 208
 QY 182 DGSFFLYSKLTVDKSRWQGNVFSCSYVMEALHNHYTQKSLSLSPGK----- 228
 Db 209 DGSFFLYSKLTVDKSRWQGNVFSCSYVMEALHNHYTQKSLSLSPGKVAAHITG 268
 QY 229 -----DW----- 230
 Db 269 LSTPPLPRGGRPKVAAHITGITRRSNSALIPISKDGLTLCQKIESWESSRKGHSFLNHV 328
 QY 231 -----LKAFYDKVAEKLKEAFDWLK-AFYDKVAEK----- 259
 Db 329 LFRNGELVIEQGLYYIYKTSYPDPPIVLMKSA 372
 QY 260 -LKEA 263
 Db 389 LMKSA 393

RESULT 14
 AAB16958
 ID AAB16958 standard; Protein; 247 AA.
 XX
 AC AAB16958;
 XX
 DT 31-OCT-2000 (first entry)
 DE FC-TMP protein sequence SEQ ID NO:6.
 XX
 KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
 KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
 KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
 KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
 KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
 KW vascular endothelial growth factor; matrix metalloproteinase;
 KW asthma; thrombosis; pharmaceutical.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO200024782-A2.
 XX
 PD 04-MAY-2000.
 XX
 PF 25-OCT-1999; 99WO-US25044.
 XX
 PR 23-OCT-1998; 98US-0105371.
 PR 22-OCT-1999; 99US-0428082.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Feige U, Liu C, Cheetham J, Boone TC;
 XX WPI; 2000-350702/30.
 DR N-PSDB; AAA69444.
 XX
 PT Novel composition of matter comprising an Fc domain and
 PT pharmacologically active peptides, useful for treating cancer and

PT autoimmune diseases -

Claim 21; Page 179-180; 608pp; English.

The present invention describes composition of matter (I) comprising an Fc domain, pharmacologically active peptides, and linkers. Where (I) is: (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2, -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4 where P1, P2, P3, and P4 = are each independently sequences of pharmacologically active peptides; L1, L2, L3, and L4 = are each independently linkers; and a, b, c, d, e, and f = are each independently 0 or 1, provided that at least 1 of a and b is 1. The composition can have cytostatic, antitumour, thrombolytic and immunosuppressive activities. DNAs, vectors and host cells from the present invention can be used for producing pharmaceutical compositions. The compositions are useful for treating cancer, asthma, thrombosis, or autoimmune diseases. The use of an Fc domain (rather than a Fab domain) can provide a longer half-life or incorporate functions such as Fc receptor binding, protein A binding, complement fixation, and possibly placental transfer. AAA69443 to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid sequences used in the exemplification of the present invention.

Sequence 247 AA;

Query Match 87.1%; Score 1246; DB 21; Length 247;
Best Local Similarity 94.3%; Pred. No. 8.8e-34;
Matches 231; Conservative 0; Mismatches 2; Indels 12; Gaps 1;

QY 1 MDRHTCPCPCAPPELLGGPSVFLPDKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
DB 1 MDRHTCPCPCAPPELLGGPSVFLPDKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
QY 61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIETKISK 120
DB 61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIETKISK 120
QY 121 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 180
DB 121 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 180
QY 181 SDGSFFLYSKLTVDKSRWQGNVFSCVMEALHNHYTQKSLSLSPGK----- 228
DB 181 SDGSFFLYSKLTVDKSRWQGNVFSCVMEALHNHYTQKSLSLSPGK----- 228
QY 229 DWLKA 233
DB 241 QWLAA 245

RESULT 15
ABB73411
ID ABB73411 standard; Protein; 247 AA.

AC ABB73411;

DT 05-APR-2002 (first entry)

DE Fc-TPO mimetic peptide (Fc-TMP) amino acid SEQ ID NO:6.

KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG;
KW EPO; erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
KW antianemic; anorectic; antiinfectivity; haemostatic; dermatological;
KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
KW sleep disorder; neurological degenerative disease; anaemia;
KW thrombocytopenia; metastatic tumour; systemic lupus erythematosus;
KW Fanconi's syndrome.

OS Homo sapiens.
OS Synthetic.

PN WO200183525-A2.

XX 08-NOV-2001.

PD 02-MAY-2001; 2001WO-US14310.

PF 03-MAY-2000; 2000US-0563286.

XX (AMGE-) AMGEN INC.

PA Felge U, Liu C, Cheetham JC, Boone TC, Gudas JM;

XX WPI; 2002-130313/17.

DR N-PSDB; ABL35761.

PT Novel vehicle-peptide molecule or its multimers useful for treating inflammatory and autoimmune diseases, cancer, rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders and infertility.

PS Claim 21; Fig 7; 176pp; English.

XX The present invention describes a vehicle-peptide molecule (I) or its multimers. (I) can have antiinflammatory, antitumour, immunosuppressive, cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological, antianemic, anorectic, antiinfectivity, haemostatic, dermatological and neuroprotective activities. (I) can be used as a therapeutic or prophylactic agent as well as for screening purposes. (I) is useful for diagnosing diseases characterised by dysfunction of their associated protein of interest, for identifying normal or abnormal proteins of interest, as a part of diagnostic kit to detect the presence of their proteins of interest in a biological sample. Additionally, (I) is useful for treating inflammatory and autoimmune diseases, tumour growth, cancer, rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders, infertility, and neurological degenerative diseases. (I), comprising EPO-mimetic compounds are useful for treating disorders characterised by low red blood cell levels such as anaemia. The TPO-mimetic comprising megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet deficiency, such as thrombocytopenia, aplastic anaemia, metastatic tumour which result in thrombocytopenia, systemic lupus erythematosus, and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777 represent amino acid and nucleic acid sequences used in the exemplification of the present invention.

XX Sequence 247 AA;

Query Match 87.1%; Score 1246; DB 23; Length 247;
Best Local Similarity 94.3%; Pred. No. 8.8e-34;
Matches 231; Conservative 0; Mismatches 2; Indels 12; Gaps 1;

QY 1 MDRHTCPCPCAPPELLGGPSVFLPDKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
DB 1 MDRHTCPCPCAPPELLGGPSVFLPDKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
QY 61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIETKISK 120
DB 61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIETKISK 120
QY 121 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 180
DB 121 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 180
QY 181 SDGSFFLYSKLTVDKSRWQGNVFSCVMEALHNHYTQKSLSLSPGK----- 228
DB 181 SDGSFFLYSKLTVDKSRWQGNVFSCVMEALHNHYTQKSLSLSPGK----- 228
QY 229 DWLKA 233
DB 241 QWLAA 245

Search completed: April 21, 2003, 10:45:28
Job time.: 33.0588 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:36 ; Search time 14.7529 Seconds
(without alignments)
526.515 Million cell updates/sec

Title: 2LINK7LINK7
Perfect score: 1430
Sequence: 1 MDKTHCTPCPCAPPELLGGPS.....AFDWLKAFYDKVAEKLKEAF 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2.6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2.6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2.6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2.6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2.6/ptodata/1/iaa/PCUS_COMB.pep.*
6: /cgn2.6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1240	86.7	389	US-09-131-247-14	Sequence 14, Appl
2	1239	86.6	660	US-09-181-706-8	Sequence 8, Appl
3	1239	86.6	660	US-09-458-791-8	Sequence 8, Appl
4	1239	86.6	660	US-09-459-066-8	Sequence 8, Appl
5	1234	86.3	347	US-07-940-861-43	Sequence 43, Appl
6	1234	86.3	347	US-08-459-512-43	Sequence 43, Appl
7	1234	86.3	347	US-08-459-657-43	Sequence 43, Appl
8	1234	86.3	347	US-08-460-132-43	Sequence 43, Appl
9	1234	86.3	347	US-08-466-465-8	Sequence 8, Appl
10	1234	86.3	347	PCT-US92-02050-43	Sequence 43, Appl
11	1234	86.3	482	US-09-189-129-2	Sequence 2, Appl
12	1233	86.2	232	US-08-595-043A-50	Sequence 50, Appl
13	1233	86.2	235	US-09-131-247-6	Sequence 6, Appl
14	1233	86.2	331	US-09-178-869-2	Sequence 2, Appl
15	1233	86.2	360	US-09-180-100-11	Sequence 11, Appl
16	1233	86.2	371	US-08-236-311-7	Sequence 7, Appl
17	1233	86.2	371	US-08-457-918-7	Sequence 7, Appl
18	1233	86.2	376	US-09-180-100-22	Sequence 22, Appl
19	1233	86.2	387	US-08-470-299-4	Sequence 4, Appl
20	1233	86.2	388	US-09-131-247-16	Sequence 16, Appl
21	1233	86.2	396	US-08-784-512-3	Sequence 3, Appl
22	1233	86.2	396	US-09-176-228-3	Sequence 3, Appl
23	1233	86.2	424	US-09-333-593A-8	Sequence 8, Appl
24	1233	86.2	424	PCT-US95-03866-12	Sequence 12, Appl
25	1233	86.2	424	PCT-US95-03866-14	Sequence 14, Appl
26	1233	86.2	437	PCT-US96-10043-11	Sequence 11, Appl
27	1233	86.2	442	PCT-US96-10043-9	Sequence 9, Appl

28	1233	86.2	446	3	US-08-397-411-7	Sequence 7, Appl
29	1233	86.2	449	1	US-08-458-516-13	Sequence 13, Appl
30	1233	86.2	459	1	US-08-157-101A-7	Sequence 7, Appl
31	1233	86.2	476	2	US-08-378-939-10	Sequence 10, Appl
32	1233	86.2	476	3	US-08-487-550-4	Sequence 4, Appl
33	1233	86.2	476	3	US-08-487-550-12	Sequence 12, Appl
34	1233	86.2	478	3	US-08-487-550-8	Sequence 8, Appl
35	1233	86.2	488	4	US-08-776-511-2	Sequence 2, Appl
36	1233	86.2	680	4	US-08-227-496C-15	Sequence 15, Appl
37	1233	86.2	711	4	US-09-485-737B-90	Sequence 90, Appl
38	1231	86.1	254	2	US-08-284-391B-33	Sequence 33, Appl
39	1231	86.1	254	4	US-09-218-950-33	Sequence 33, Appl
40	1228	85.9	664	3	US-08-957-063-16	Sequence 16, Appl
41	1228	85.9	664	3	US-08-957-063-18	Sequence 18, Appl
42	1228	85.9	664	4	US-09-487-685-16	Sequence 16, Appl
43	1228	85.9	664	4	US-09-487-685-18	Sequence 18, Appl
44	1228	85.9	664	4	US-08-802-805D-16	Sequence 16, Appl
45	1228	85.9	664	4	US-08-802-805D-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-09-131-247-14
; Sequence 14, Application US/09131247
; Patent No. 6294170
; GENERAL INFORMATION:
; APPLICANT: Boone, Thomas C.
; APPLICANT: Herhenson, Susan
; APPLICANT: Bevilacqua, Michael P.
; APPLICANT: Collins, David S.
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR TREATING INFLAMMATORY
; FILE REFERENCE: A-365F
; CURRENT APPLICATION NUMBER: US/09/131,247
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: 60/055,185
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: PCT/US 97/02131
; EARLIER FILING DATE: 1997-02-10
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 14
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Human
US-09-131-247-14

Query Match 86.7%; Score 1240; DB 4; Length 389;
Best Local Similarity 63.7%; Pred. No. 8.4e-40;
Matches 239; Conservative 5; Mismatches 8; Indels 123; Gaps 7;

QY	2	DKTHCTPCAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD	61
DB	10	DKTHCTPCAPPELLGGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD	69
QY	62	GVEVHNAKTPREEQNSTYRVVSVLTVLDHDLNGKEYCKVSNKALPAPIEKTISKAK	121
DB	70	GVEVHNAKTPREEQNSTYRVVSVLTVLDHDLNGKEYCKVSNKALPAPIEKTISKAK	129
QY	122	GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD	181
DB	130	GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD	189
QY	182	DGSFFLYSKLTVDKSRWQQGNVSCVMHEALNHHYTKQSLSLSPGK	228
DB	190	DGSFFLYSKLTVDKSRWQQGNVSCVMHEALNHHYTKQSLSLSPGKMRPSGRKSKMQA	249
QY	229	--DW---LKAFY-----DKV-----	238
DB	250	FRIDVNDQKTFYLRNNOLVAGYLQGNVLEEKIDVVPFHALFGLIHGGMCLSCVKS	309

Qy	2	DKTHTCPCPCAPBELLGGPSVFLFPPKPKD	TLMI	STPEVTCVVVD	SHDEDPVKFNK	XVD	61
Db	37	DKTHTCPCPCAPBEGAGPSVFLFPPKPKD	TLMI	STPEVTCVVVD	SHDEDPVKFNK	XVD	96
Qy	62	GVEVHNAKTPREEQYNSTYRVVSVLTVL	QHQL	WLGNGKEYCK	RVSNKALP	APIETKISAK	121
Db	97	GVEVHNAKTPREEQYNSTYRVVSVLTVL	QHQL	WLGNGKEYCK	RVSNKALP	APIETKISAK	156
Qy	122	GQPREPQVYTLPPSRDELTKNQVSLTCL	VKG	FPYSDIAVEWES	NGQPENNY	KTTTPPV	181

Query Match	86.68;	Score 1239;	DB 4;	Length 660;
-------------	--------	-------------	-------	-------------

```

; TELEPHONE: (206)470-4189
; TELEFAX: (206)233-0644
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 660 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-459-066-8

Query Match 86.6%; Score 1239; DB 4; Length 660;
Best Local Similarity 48.5%; Pred. No. 3.3e-39;
Matches 238; Conservative 7; Mismatches 18; Indels 228; Gaps 6;

QY 2 DKHTCCPCPAPELLGGPSVFLFPPPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 37 DKHTCCPCPAPEAGAPSFLFPPPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 96
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 62 GVEVHNKATPREEOYNSTRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 97 GVEVHNKATPREEOYNSTRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 156
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPV 181
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 157 GQPREPQVYTLPPSRDEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPV 216
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 182 DGSFFLYSKITVDKSRWQQGNVSCSYMHEALHNHYTQKSLSLSPGK----- 228
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 217 DGSFFLYSKITVDKSRWQQGNVSCSYMHEALHNHYTQKSLSLSPGKGGGGGGSTSI 276
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 229 DWLK----- 232
Db : | |
QY 277 EHWKFEISSELIITYLLIDDDVLYTGVCNAVYFSSNNELNKTGLTNNNNYITSIKVEDTLV 336
QY 233 ----- 232
Db
QY 337 CGTNNGPKCWKIDGSEDPKYRGRGYAPQNSKVTIISHNECVLSDINISKEGIKRWRRF 396

```

```

Db 397 DPGCYDLYTADNVIPKDGVRGAFVDKGYDKVYLFTDTIDTKRIVKIPYIAQMCLND 456
Qy 248 -----W-----LKFADKVAEK-----259
Db 457 EGGPSSLSSHRWSTFLKAVELECDIGRSYRQIHSKAKITDNDTILYVFFDSPYSKALC 516
Qy 260 -----LKEAF 264
Db 517 TYSMNAIKHSF 527

RESULT 5
US-07-940-861-43
; Sequence 43, Application US/07940861
; Patent No. 5547853
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: WALLNER, Barbara P.
; APPLICANT: MILLER, Glenn T.
; APPLICANT: ROSA, Margaret D.
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
; TITLE OF INVENTION: FUNCTION ASSOCIATED ANTIGEN 3
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 875 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022-6250
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/940,861
; FILING DATE: 21-OCT-1992
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02050
; FILING DATE: 12-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/667,971
; FILING DATE: 12-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,967
; FILING DATE: 07-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY, James F., Jr.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B151CIP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)715-0600
; TELEFAX: (212)715-0673
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-940-861-43

```

```

Query Match      86.3%; Score 1234; DB 1; Length 347;
Best Local Similarity 99.6%; Pred. No. 1.1e-39;
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHCPCPAPPELLGGSVLFPKPKDITLMISRTPEVTCVVDVSHEDPEVKFNWY 50
Db :|||||
Db 120 VDKTHCPCPAPPELLGGSVLFPKPKDITLMISRTPEVTCVVDVSHEDPEVKFNWY 179
QY 61 DGVEVHNATKPREQYNSTYRVSVLTVLHODWLNKGEYCKVSNKALPAPIETKISK 120
Db :|||||
Db 180 DGVEVHNATKPREQYNSTYRVSVLTVLHODWLNKGEYCKVSNKALPAPIETKISK 239
QY 121 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 180
Db :|||||
Db 240 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 299
QY 181 SDGSFELYSLKLTVDKSRWQGNVFCSCVMHEALHNNHYTOKSLSPGK 228
Db :|||||
Db 300 SDGSFELYSLKLTVDKSRWQGNVFCSCVMHEALHNNHYTOKSLSPGK 347

```

```

RESULT 6
US-08-459-512-43
; Sequence 43, Application US/08459512
; Patent No. 5728677
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: WALLNER, Barbara P.
; APPLICANT: MILLER, Glenn T.
; APPLICANT: ROSA, Margaret D.
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022-6250
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,512
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02050
; FILING DATE: 12-MAR-1992
; APPLICATION NUMBER: US 07/667,971
; FILING DATE: 12-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,967
; FILING DATE: 07-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY, James F., Jr.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B151CIP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)715-0600
; TELEFAX: (212)715-0673
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-459-512-43

```

```

Query Match      86.3%; Score 1234; DB 1; Length 347;
Best Local Similarity 99.6%; Pred. No. 1.1e-39;
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHCPCPAPPELLGGSVLFPKPKDITLMISRTPEVTCVVDVSHEDPEVKFNWY 50
Db :|||||
Db 120 VDKTHCPCPAPPELLGGSVLFPKPKDITLMISRTPEVTCVVDVSHEDPEVKFNWY 179
QY 61 DGVEVHNATKPREQYNSTYRVSVLTVLHODWLNKGEYCKVSNKALPAPIETKISK 120
Db :|||||
Db 180 DGVEVHNATKPREQYNSTYRVSVLTVLHODWLNKGEYCKVSNKALPAPIETKISK 239
QY 121 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 180
Db :|||||
Db 240 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 299
QY 181 SDGSFELYSLKLTVDKSRWQGNVFCSCVMHEALHNNHYTOKSLSPGK 228
Db :|||||
Db 300 SDGSFELYSLKLTVDKSRWQGNVFCSCVMHEALHNNHYTOKSLSPGK 347

```

```

RESULT 7
US-08-459-657-43
; Sequence 43, Application US/08459657
; Patent No. 5914111
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: WALLNER, Barbara P.
; APPLICANT: MILLER, Glenn T.
; APPLICANT: ROSA, Margaret D.
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022-6250
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,657
FILING DATE: 02-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/02050
FILING DATE: 12-MAR-1992
APPLICATION NUMBER: US 07/667,971
FILING DATE: 12-MAR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/770,967
FILING DATE: 07-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: HALEY, James F., Jr.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B151CIP2
TELEPHONE: (212)715-0600
TELEFAX: (212)715-0673
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 347 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-459-657-43

Query Match 86.3%; Score 1234; DB 2; Length 347;
Best Local Similarity 99.6%; Pred. No. 1.1e-39;
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHCTPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60
Db 120 VDKTHCTPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 179
QY 61 DGEVHNATKPREQYNSTYRVSVLTVLHQDLNGKEYKCKVSNKALPAPIEKTISKA 120
Db 180 DGEVHNATKPREQYNSTYRVSVLTVLHQDLNGKEYKCKVSNKALPAPIEKTISKA 239
QY 121 KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYFPSDIAVWESNGQPENNYKTTPPVLD 180
Db 240 KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYFPSDIAVWESNGQPENNYKTTPPVLD 299
QY 181 SDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 228
Db 300 SDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 347

RESULT 8
US-08-460-132-43
Sequence 43, Application US/08460132
Patent No. 5928643
GENERAL INFORMATION:
APPLICANT: BIOGEN, INC.
APPLICANT: WALLNER, Barbara P.
APPLICANT: MILLER, Glenn T.
APPLICANT: ROSA, Margaret D.
TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Neave
STREET: 875 Third Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10022-6250
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,132
FILING DATE: 02-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/940,861
FILING DATE: 21-OCT-1992
APPLICATION NUMBER: PCT/US92/02050
FILING DATE: 12-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/667,971
FILING DATE: 12-MAR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/770,967
FILING DATE: 07-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: HALEY, James F., Jr.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B151CIP2
TELEPHONE: (212)715-0600
TELEFAX: (212)715-0673
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 347 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-460-132-43

Query Match 86.3%; Score 1234; DB 2; Length 347;
Best Local Similarity 99.6%; Pred. No. 1.1e-39;
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHCTPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60
Db 120 VDKTHCTPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 179
QY 61 DGEVHNATKPREQYNSTYRVSVLTVLHQDLNGKEYKCKVSNKALPAPIEKTISKA 120
Db 180 DGEVHNATKPREQYNSTYRVSVLTVLHQDLNGKEYKCKVSNKALPAPIEKTISKA 239
QY 121 KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYFPSDIAVWESNGQPENNYKTTPPVLD 180
Db 240 KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYFPSDIAVWESNGQPENNYKTTPPVLD 299
QY 181 SDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 228
Db 300 SDGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 347

RESULT 9
US-08-466-465-8
Sequence 8, Application US/08466465
Patent No. 6162432
GENERAL INFORMATION:
APPLICANT: Wallner, Barbara P.
APPLICANT: Cooper, Kevin D.
TITLE OF INVENTION: Method of Prophylaxis or Treatment of Antigen
PRESENTING CELL DRIVEN SKIN CONDITIONS USING
INHIBITORS OF THE CD2/LFA-3 INTERACTION
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,465
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08755
; FILING DATE: 06-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/862,022
; FILING DATE: 12-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,969
; FILING DATE: 07-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis (PLM)
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGP-111CP
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-466-465-8

Query Match      86.3%; Score 1234; DB 4; Length 347;
Best Local Similarity 99.6%; Pred. No. 1.1e-39;
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHCTPCPAPPELLGSPSVFLPPPKDFTLMISRTPEVTCVVVDVSHEDPEVKFNMYV 60
   :|||||
Db 120 VDKTHCTPCPAPPELLGSPSVFLPPPKDFTLMISRTPEVTCVVVDVSHEDPEVKFNMYV 179

QY 61 DGEVHNNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
   :|||||
Db 180 DGEVHNNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 239

QY 121 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 180
   :|||||
Db 240 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 299

QY 181 SDGSFFLYSKLTVDSKRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 228
   :|||||
Db 300 SDGSFFLYSKLTVDSKRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 347

RESULT 10
PCT-US92-02050-43
; Sequence 43, Application PC/TUS9202050
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: WALLNER, Barbara P.
; APPLICANT: MILLER, Glenn T.
; APPLICANT: ROSA, Margaret D.
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
; TITLE OF INVENTION: FUNCTION ASSOCIATED ANTIGEN 3
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 875 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022-6250
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02050
; FILING DATE: 19920312
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/667,971
; FILING DATE: 12-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,967
; FILING DATE: 07-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY, James F., Jr.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B15C1P2
; TELEPHONE: (212)715-0600
; TELEFAX: (212)715-0673
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US92-02050-43

Query Match      86.3%; Score 1234; DB 5; Length 347;
Best Local Similarity 99.6%; Pred. No. 1.1e-39;
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHCTPCPAPPELLGSPSVFLPPPKDFTLMISRTPEVTCVVVDVSHEDPEVKFNMYV 60
   :|||||
Db 120 VDKTHCTPCPAPPELLGSPSVFLPPPKDFTLMISRTPEVTCVVVDVSHEDPEVKFNMYV 179

QY 61 DGEVHNNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
   :|||||
Db 180 DGEVHNNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 239

QY 121 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 180
   :|||||
Db 240 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 299

QY 181 SDGSFFLYSKLTVDSKRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 228
   :|||||
Db 300 SDGSFFLYSKLTVDSKRWQGNVFCSCVMHEALHNHYTQKSLSLSPGK 347

RESULT 11
US-09-189-129-2
; Sequence 2, Application US/09189129
; Patent No. 6323027
; GENERAL INFORMATION:
; APPLICANT: Burkly, Linda C
; APPLICANT: Benjamin, Christopher D
; APPLICANT: Hession, Catherine A
; APPLICANT: Whitty, Adrian
; TITLE OF INVENTION: COMMON GAMMA CHAIN BLOCKING AGENTS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Biogen, Inc.
; STREET: 14 Cambridge Center
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02142
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/189,129
;; FILING DATE:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: A006 PCT CIP
;; FILING DATE: 09-MAY-1997
;; APPLICATION NUMBER: 60/017,466
;; FILING DATE: 10-MAY-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Kaplan, Warren A.
;; REGISTRATION NUMBER: 34,199
;; REFERENCE/DOCKET NUMBER: A006 PCT CIP
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 617 679-2000
;; TELEFAX: 617 679-2838
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 482 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: YES
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
US-09-189-129-2

Query Match 86.3%; Score 1234; DB 4; Length 482;
Best Local Similarity 99.6%; Pred. No. 2.4e-39;
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDKTHCTPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60
Db 255 VDKTHCTPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 314
QY 61 DGEVHNKTKPREEQYNSTYRVSVTLVLDHQLNGKEYCKVSKNKPAPIEKTISKAK 120
Db 315 DGEVHNKTKPREEQYNSTYRVSVTLVLDHQLNGKEYCKVSKNKPAPIEKTISKAK 374
QY 121 KGPREQVYTLPPSRDELTKNOVSLTCLVKGYFSPDIAVEWESNGQPENNYKTTPPVLD 180
Db 375 KGPREQVYTLPPSRDELTKNOVSLTCLVKGYFSPDIAVEWESNGQPENNYKTTPPVLD 434
QY 181 SDGSFFLYSKLTVDKSRWQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
Db 435 SDGSFFLYSKLTVDKSRWQGNVFSCSVMHEALHNHYTQKSLSLSPGK 482

RESULT 12
US-08-595-043A-50
Sequence 50, Application US/08595043A
Patent No. 5935824
GENERAL INFORMATION:
APPLICANT: SGARLATO, GREGORY D.
TITLE OF INVENTION: PROTEIN EXPRESSION SYSTEM
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: MEDLEN & CARROLL
STREET: 220 MONTGOMERY STREET, SUITE 2200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/595,043A
FILING DATE: 31-JAN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CARROLL, PETER G.

;; REGISTRATION NUMBER: 32,837
;; REFERENCE/DOCKET NUMBER: SGAR-00371
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 705-8410
;; TELEFAX: (415) 397-8338
;; INFORMATION FOR SEQ ID NO: 50:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 232 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-595-043A-50

Query Match 86.2%; Score 1233; DB 2; Length 232;
Best Local Similarity 100.0%; Pred. No. 4.5e-40;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 DKTHCTPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61
Db 6 DKTHCTPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 65
QY 62 GVEVHNKTKPREEQYNSTYRVSVTLVLDHQLNGKEYCKVSKNKPAPIEKTISKAK 121
Db 66 GVEVHNKTKPREEQYNSTYRVSVTLVLDHQLNGKEYCKVSKNKPAPIEKTISKAK 125
QY 122 GOPREQVYTLPPSRDELTKNOVSLTCLVKGYFSPDIAVEWESNGQPENNYKTTPPVLD 181
Db 126 GOPREQVYTLPPSRDELTKNOVSLTCLVKGYFSPDIAVEWESNGQPENNYKTTPPVLD 185
QY 182 DGSFFLYSKLTVDKSRWQGNVFSCSVMHEALHNHYTQKSLSLSPGK 228
Db 186 DGSFFLYSKLTVDKSRWQGNVFSCSVMHEALHNHYTQKSLSLSPGK 232

RESULT 13
US-09-131-247-6
Sequence 6, Application US/09131247
Patent No. 6294170
GENERAL INFORMATION:
APPLICANT: Boone, Thomas C.
APPLICANT: Herhenson, Susan
APPLICANT: Bevilacqua, Michael P.
APPLICANT: Collins, David S.
TITLE OF INVENTION: COMPOSITION AND METHOD FOR TREATING INFLAMMATORY
DISEASES
FILE REFERENCE: A-365F.
CURRENT APPLICATION NUMBER: US/09/131,247
CURRENT FILING DATE: 1998-08-07
EARLIER APPLICATION NUMBER: 60/055,185
EARLIER FILING DATE: 1997-08-08
EARLIER APPLICATION NUMBER: PCT/US 97/02131
EARLIER FILING DATE: 1997-02-10
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 235
TYPE: PRT
ORGANISM: Human
US-09-131-247-6

Query Match 86.2%; Score 1233; DB 4; Length 235;
Best Local Similarity 100.0%; Pred. No. 4.7e-40;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 DKTHCTPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61
Db 9 DKTHCTPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 68
QY 62 GVEVHNKTKPREEQYNSTYRVSVTLVLDHQLNGKEYCKVSKNKPAPIEKTISKAK 121
Db 69 GVEVHNKTKPREEQYNSTYRVSVTLVLDHQLNGKEYCKVSKNKPAPIEKTISKAK 128
QY 122 GOPREQVYTLPPSRDELTKNOVSLTCLVKGYFSPDIAVEWESNGQPENNYKTTPPVLD 181

Db 129 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 188
QY 182 DGSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTOKLSLSPGK 228
Db 189 DGSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTOKLSLSPGK 235
RESULT 14
US-09-178-869-2
; Sequence 2, Application US/09178869B
; Patent No. 6197294
; GENERAL INFORMATION:
; APPLICANT: Wong, Shou
; APPLICANT: Hickey, William F.
; APPLICANT: Hamman, Joseph P.
; APPLICANT: Baetge, E. Edward
; TITLE OF INVENTION: CELL SURFACE-INDUCED MACROPHAGE ACTIVATION
; FILE REFERENCE: 17810-043
; CURRENT APPLICATION NUMBER: US/09/178,869B
; CURRENT FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-178-869-2

Query Match 86.2%; Score 1233; DB 4; Length 331;
Best Local Similarity 100.0%; Pred. No. 1.1e-39;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 DKTHCTCPCPAPPELLGGPSVFLFPPPKDRLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
Db 105 DKTHCTCPCPAPPELLGGPSVFLFPPPKDRLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 164
QY 62 GVEVHNATKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
Db 165 GVEVHNATKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 224
QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 181
Db 225 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 284
QY 182 DGSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTOKLSLSPGK 228
Db 285 DGSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTOKLSLSPGK 331

RESULT 15
US-09-180-100-11
; Sequence 11, Application US/09180100
; Patent No. 6306395
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, No. 630639510
; APPLICANT: NAGATA, Shigekazu
; TITLE OF INVENTION: NOVEL FAS ANTIGEN DERIVATIVE
; FILE REFERENCE: 1110-207P
; CURRENT APPLICATION NUMBER: US/09/180,100
; CURRENT FILING DATE: 1998-11-02
; EARLIER APPLICATION NUMBER: PCT/JP97/01502
; EARLIER FILING DATE: 1997-05-01
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-180-100-11

Query Match 86.2%; Score 1233; DB 4; Length 360;

Best Local Similarity 100.0%; Pred. No. 1.3e-39;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 DKTHCTCPCPAPPELLGGPSVFLFPPPKDRLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
Db 134 DKTHCTCPCPAPPELLGGPSVFLFPPPKDRLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 193
QY 62 GVEVHNATKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
Db 194 GVEVHNATKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 253
QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 181
Db 254 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 313
QY 182 DGSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTOKLSLSPGK 228
Db 314 DGSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTOKLSLSPGK 360

Search completed: April 21, 2003, 10:50:27
Job time : 16.7529 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:43:21 ; Search time 12.1647 Seconds
(without alignments)
1640.982 Million cell updates/sec

Title: 2LINK7LINK7
Perfect score: 1430
Sequence: 1 MDKHTCPPCPAPPELLGGPS.....AFDNLKAFYDKVAEKLKEAF 264

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.0

Searched: 288829 seqs, 7561385 residues
Total number of hits satisfying chosen parameters: 288829

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published_Applications_AA:*
- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1260	88.1	729	10	US-09-825-012-52
2	1260	88.1	739	10	US-09-825-012-61
3	1257	87.9	731	10	US-09-825-012-46
4	1257	87.9	741	10	US-09-825-012-55
5	1250	87.4	730	10	US-09-825-012-49
6	1250	87.4	740	10	US-09-825-012-58
7	1240	86.7	389	10	US-09-784-623-14
8	1238	86.6	228	9	US-09-840-277-2
9	1238	86.6	228	9	US-09-847-249A-2
10	1238	86.6	228	9	US-09-840-669B-2
11	1238	86.6	228	9	US-09-843-221A-2
12	1238	86.6	228	10	US-09-847-712-2
13	1238	86.6	282	9	US-09-840-277-109
14	1237	86.5	394	10	US-09-854-864-31
15	1234	86.3	229	9	US-10-215-297-2
16	1234	86.3	229	9	US-10-215-298-2
17	1234	86.3	347	9	US-10-091-236-17
18	1234	86.3	347	9	US-10-091-313-7
19	1234	86.3	347	9	US-10-091-268-7

20	1234	86.3	347	10	US-09-796-033-8	Sequence 8, Appli
21	1234	86.3	347	10	US-09-730-465-8	Sequence 8, Appli
22	1234	86.3	380	10	US-09-948-018-36	Sequence 36, Appli
23	1234	86.3	380	10	US-09-948-018-39	Sequence 39, Appli
24	1234	86.3	388	10	US-09-734-300-8	Sequence 8, Appli
25	1234	86.3	388	10	US-09-734-300-9	Sequence 9, Appli
26	1234	86.3	399	9	US-09-832-659-2	Sequence 2, Appli
27	1234	86.3	418	9	US-09-832-659-42	Sequence 42, Appli
28	1234	86.3	423	9	US-09-832-659-44	Sequence 44, Appli
29	1234	86.3	482	10	US-09-824-286-2	Sequence 2, Appli
30	1234	86.3	594	10	US-09-815-108-22	Sequence 22, Appli
31	1233	86.2	232	10	US-09-996-357-10	Sequence 10, Appli
32	1233	86.2	235	10	US-09-784-623-6	Sequence 6, Appli
33	1233	86.2	247	10	US-09-996-357-13	Sequence 13, Appli
34	1233	86.2	267	10	US-09-996-357-12	Sequence 12, Appli
35	1233	86.2	281	10	US-09-854-864-10	Sequence 10, Appli
36	1233	86.2	283	10	US-09-854-864-9	Sequence 9, Appli
37	1233	86.2	330	9	US-10-047-542-20	Sequence 20, Appli
38	1233	86.2	330	9	US-09-995-898A-15	Sequence 15, Appli
39	1233	86.2	360	10	US-09-949-713-11	Sequence 11, Appli
40	1233	86.2	376	10	US-09-949-713-22	Sequence 22, Appli
41	1233	86.2	388	10	US-09-784-623-16	Sequence 16, Appli
42	1233	86.2	397	10	US-09-854-864-18	Sequence 18, Appli
43	1233	86.2	404	10	US-09-948-018-16	Sequence 16, Appli
44	1233	86.2	443	9	US-10-166-232A-5	Sequence 5, Appli
45	1233	86.2	451	9	US-10-077-023-17	Sequence 17, Appli

ALIGNMENTS

RESULT 1
US-09-825-012-52
; Sequence 52, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 52
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFG1 heavy chain - DNase I fusion
US-09-825-012-52

Query Match	88.1%	Score	1260;	DB	10;	Length	729;
Best Local Similarity	50.5%	Pred	No. 1.6e-31;				
Matches	241;	Conservative	6;	Mismatches	15;	Indels	215;
Gaps	5;						
QY	2	DKHTCTPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNMYVD	61				
Db	241	DKHTCTPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNMYVD	300				
QY	62	GVEVHNKATPREQYNTYRVSVLTVLHQDWLNGKEYKCKVSNKALPAPIETISKAK	121				
Db	301	GVEVHNKATPREQYNTYRVSVLTVLHQDWLNGKEYKCKVSNKALPAPIETISKAK	360				
QY	122	GQPEPQVYTLPPSRDELTKNOVSLTCLVKGFPYSDAVESNGQPNNTKTPPVLDLS	181				
Db	361	GQPEPQVYTLPPSRDELTKNOVSLTCLVKGFPYSDAVESNGQPNNTKTPPVLDLS	420				
QY	182	DGSPFLYSLKLTVDKRSWQOGNVSFSCVSHAEALHNHYTOKLSLSLSPGKDWLK-----	232				

Db 421 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGSGGLKIAAFNIQTF 480
QY 233 -----KAFYDKV-----AKLKEA 263
Db 481 GETKMSNATLVSYIOLSRDYALVQEVDRDLSHTAVGKLLDNLNODAPDTHYHYVSEPL 540
QY 233 -----KAFYDKV-----AKLKEA 263
Db 541 GRNSYKERYLVYRPPDOVSADSYDDGCEPCGNDTFNREPAIVRFFSRETEVREFAIV 600
QY 233 -----KAFYDKV-----AKLKEA 263
Db 601 PLHAAPGDAVAEIDALYDVLVDQEKWGLEDMVLMGDFNAGCSYVRPSQWSSIRLWTSPT 660
QY 246 FOWL-----KAFYDKV-----AKLKEA 263
Db 661 FOWLIPDSADTTATPTHCAYDRIVVAGMLLRGAVPDSALPFNFQAAYGLSDOLAQA 717
RESULT 2
US-09-825-012-61
; Sequence 61, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; LENGTH: 739
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFG1 heavy chain - DNase I fusion
US-09-825-012-61
Query Match 88.1%; Score 1260; DB 10; Length 739;
Best Local Similarity 50.5%; Pred. No. 1.6e-31;
Matches 241; Conservative 6; Mismatches 15; Indels 215; Gaps 5;
QY 2 DKHTCCPCPAPPELLGGPSVFLFPPPKDRLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
Db 241 DKHTCCPCPAPPELLGGPSVFLFPPPKDRLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 300
QY 62 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSKNALKALPAPIETISKAK 121
Db 301 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSKNALKALPAPIETISKAK 360
QY 122 GOPREPVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 181
Db 361 GOPREPVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 420
QY 182 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGSGGLKIAAFNIQTF 480
Db 421 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGSGGLKIAAFNIQTF 540
QY 233 -----KAFYDKV-----AKLKEA 263
Db 481 GETKMSNATLVSYIOLSRDYALVQEVDRDLSHTAVGKLLDNLNODAPDTHYHYVSEPL 540
QY 233 -----KAFYDKV-----AKLKEA 263
Db 541 GRNSYKERYLVYRPPDOVSADSYDDGCEPCGNDTFNREPAIVRFFSRETEVREFAIV 600
QY 233 -----KAFYDKV-----AKLKEA 263

Db 601 PLHAAPGDAVAEIDALYDVLVDQEKWGLEDMVLMGDFNAGCSYVRPSQWSSIRLWTSPT 660
QY 246 FOWL-----KAFYDKV-----AKLKEA 263
Db 661 FOWLIPDSADTTATPTHCAYDRIVVAGMLLRGAVPDSALPFNFQAAYGLSDOLAQA 717
RESULT 3
US-09-825-012-46
; Sequence 46, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
; LENGTH: 731
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFG1 heavy chain - DNase I fusion
US-09-825-012-46
Query Match 87.9%; Score 1257; DB 10; Length 731;
Best Local Similarity 50.1%; Pred. No. 2e-31;
Matches 240; Conservative 6; Mismatches 16; Indels 217; Gaps 4;
QY 2 DKHTCCPCPAPPELLGGPSVFLFPPPKDRLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
Db 241 DKHTCCPCPAPPELLGGPSVFLFPPPKDRLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 300
QY 62 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSKNALKALPAPIETISKAK 121
Db 301 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSKNALKALPAPIETISKAK 360
QY 122 GOPREPVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 181
Db 361 GOPREPVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 420
QY 182 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGSGGLKIAAFNIQ 480
Db 421 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSPGSGGLKIAAFNIQ 540
QY 229 -----KAFYDKV-----AKLKEA 263
Db 481 GETKMSNATLVSYIOLSRDYALVQEVDRDLSHTAVGKLLDNLNODAPDTHYHYVSE 540
QY 229 -----KAFYDKV-----AKLKEA 263
Db 541 GRNSYKERYLVYRPPDOVSADSYDDGCEPCGNDTFNREPAIVRFFSRETEVREFA 600
QY 229 -----KAFYDKV-----AKLKEA 263
Db 601 PLHAAPGDAVAEIDALYDVLVDQEKWGLEDMVLMGDFNAGCSYVRPSQWSSIRLWTS 660
QY 244 FOWL-----KAFYDKV-----AKLKEA 263
Db 661 FOWLIPDSADTTATPTHCAYDRIVVAGMLLRGAVPDSALPFNFQAAYGLSDOLAQA 717
RESULT 4
US-09-825-012-55
; Sequence 55, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:

```
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55
; LENGTH: 741
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFG1 heavy chain - DNase I fusion
US-09-825-012-55

Query Match      87.9%; Score 1257; DB 10; Length 741;
Best Local Similarity 50.1%; Pred. No. 2e-31;
Matches 240; Conservative 6; Mismatches 16; Indels 217; Gaps 4;

QY 2 DKHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61
Db 241 DKHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 300
QY 62 GVEVHNKTKPREEQYNSTYRVVSVLTVLHODWLNQGEYCKCKVSNKALPAPIETISKAK 121
Db 301 GVEVHNKTKPREEQYNSTYRVVSVLTVLHODWLNQGEYCKCKVSNKALPAPIETISKAK 360
QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 181
Db 361 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 420
QY 182 DGSFPLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK----- 228
Db 421 DGSFPLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK----- 480
QY 229 -----KAFYDKV-----AEKLKEA 263
Db 481 TFGTKMSNATLVSYIVQILSRDYALVQEVDRSHLTAVGKLLDNLDNADPTTHYVVSE 540
QY 229 ----- 228
Db 541 PLGRNSYKERYLFVYRPDOVSADSVYDDGCEPCGNDTFNREPAIVRFFSRTFVREFA 600
QY 229 -----DWLKAFYDKVAEK-----LK 243
Db 601 IVPLHAAPGDAVAEIDALYDVLVDQEKWGLDVMGMDFNAGCSYVRPQWSSIRLWTS 660
QY 244 EAFDWL-----KAFYDKV-----AEKLKEA 263
Db 661 PTFOWLIPDSADTTATPTTHCAIDRVVAGMLLRGAVVPDSALPFFNFOAAYGLSDQLAQA 719

RESULT 5
US-09-825-012-49
; Sequence 49, Application US/09825012
; Patent No. US200201227981
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49

Query Match      87.9%; Score 1257; DB 10; Length 740;
Best Local Similarity 50.4%; Pred. No. 3.4e-31;
Matches 241; Conservative 6; Mismatches 15; Indels 216; Gaps 6;

QY 2 DKHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61
Db 241 DKHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 300
QY 62 GVEVHNKTKPREEQYNSTYRVVSVLTVLHODWLNQGEYCKCKVSNKALPAPIETISKAK 121
Db 301 GVEVHNKTKPREEQYNSTYRVVSVLTVLHODWLNQGEYCKCKVSNKALPAPIETISKAK 360
QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 181
Db 361 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 420
QY 182 DGSFPLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK----- 232
Db 421 DGSFPLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK----- 480
QY 233 ----- 232
Db 481 FGETKMSNATLVSYIVQILSRDYALVQEVDRSHLTAVGKLLDNLDNADPTTHYVVSE 540
QY 233 ----- 232
Db 541 LGRNSYKERYLFVYRPDOVSADSVYDDGCEPCGNDTFNREPAIVRFFSRTFVREFAI 600
QY 233 -----KVAEK-----LKE 244
Db 601 VPLHAAPGDAVAEIDALYDVLVDQEKWGLDVMGMDFNAGCSYVRPQWSSIRLWTS 660
QY 245 AFDWL-----KAFYDKV-----AEKLKEA 263
Db 661 TFWLIPDSADTTATPTTHCAIDRVVAGMLLRGAVVPDSALPFFNFOAAYGLSDQLAQA 718

RESULT 6
US-09-825-012-58
; Sequence 58, Application US/09825012
; Patent No. US200201227981
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58
; LENGTH: 740
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFG1 heavy chain - DNase I fusion
US-09-825-012-58

Query Match      87.4%; Score 1250; DB 10; Length 730;
Best Local Similarity 50.4%; Pred. No. 3.3e-31;
Matches 241; Conservative 6; Mismatches 15; Indels 216; Gaps 6;

QY 2 DKHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61
Db 241 DKHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 300
QY 62 GVEVHNKTKPREEQYNSTYRVVSVLTVLHODWLNQGEYCKCKVSNKALPAPIETISKAK 121
Db 301 GVEVHNKTKPREEQYNSTYRVVSVLTVLHODWLNQGEYCKCKVSNKALPAPIETISKAK 360
QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 181
Db 361 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 420
QY 182 DGSFPLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK----- 232
Db 421 DGSFPLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK----- 480
QY 233 ----- 232
Db 481 FGETKMSNATLVSYIVQILSRDYALVQEVDRSHLTAVGKLLDNLDNADPTTHYVVSE 540
QY 233 ----- 232
Db 541 LGRNSYKERYLFVYRPDOVSADSVYDDGCEPCGNDTFNREPAIVRFFSRTFVREFAI 600
QY 233 -----AFYD---KVAEK-----LKE 244
Db 601 VPLHAAPGDAVAEIDALYDVLVDQEKWGLDVMGMDFNAGCSYVRPQWSSIRLWTS 660
QY 245 AFDWL-----KAFYDKV-----AEKLKEA 263
Db 661 TFWLIPDSADTTATPTTHCAIDRVVAGMLLRGAVVPDSALPFFNFOAAYGLSDQLAQA 718
```

```
Db 241 DKHTCTPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYD 300
QY 62 GVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 121
Db 301 GVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 360
QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 181
Db 361 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 420
QY 182 DGSFFLYSKLTVDKSRWQGNVFSCVMHEALHNHYTQKSLSLSP- GKDWL- 232
Db 421 DGSFFLYSKLTVDKSRWQGNVFSCVMHEALHNHYTQKSLSLSPGSGLKIAAFNIQT 480
QY 233 - 232
Db 481 FGETKMSNATLVSYI VOILSRDYDIALVQEVRODLSHTAVGKLLDNLNODAPDTYHYVYVSEP 540
QY 233 - 232
Db 541 LGRNSYKERYLVYRPDQVSAVDYSDYDDGCEPCGNDTFNREPAIYVRFSTRTEVREFAI 600
QY 233 - -AFYD---KVAEK-----LKE 244
Db 601 VPLHAAPGDAVAEIDALYDLVDQEKWGLDYMVGDFNAGCSYVVRPQWSSIRLWTSP 660
QY 245 AFDWL-----KAFYDKV-----AEKLEA 263
Db 661 TFOWLIPDSADTTATPTTHCAIDRIVVAGMLLRGAVVPSALPFOAAYGLSDQLAQ 718
RESULT 7
US-09-784-623-14
; Sequence 14, Application US/09784623
; Patent No. US20020009454A1
; GENERAL INFORMATION:
; APPLICANT: Boone, Thomas C.
; APPLICANT: Hershenson, Susan
; APPLICANT: Bevilacqua, Michael P.
; APPLICANT: Collins, David S.
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR TREATING INFLAMMATORY
; DISEASES
; FILE REFERENCE: A-365F
; CURRENT APPLICATION NUMBER: US/09/784,623
; CURRENT FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: 09/131,247
; PRIOR FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: PCT/US 97/02131
; PRIOR FILING DATE: 1997-02-10
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Human
US-09-784-623-14
Query Match 86.7%; Score 1240; DB 10; Length 389;
Best Local Similarity 63.7%; Pred. No. 1.6e-31;
Matches 239; Conservative 5; Mismatches 8; Indels 123; Gaps 7;
QY 2 DKHTCTPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYD 61
Db 10 DKHTCTPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYD 69
QY 62 GVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 121
Db 70 GVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKAK 129
QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 181
Db 130 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 189
```

```
QY 182 DGSFFLYSKLTVDKSRWQGNVFSCVMHEALHNHYTQKSLSLSPGK----- 238
Db 190 DGSFFLYSKLTVDKSRWQGNVFSCVMHEALHNHYTQKSLSLSPGKMRPSGRKSKMQA 249
QY 229 --DW---LKAFY-----DKV----- 238
Db 250 FRWDVYNQKTFYLRNQLVAGYVLOGPNVLEEKIDVVPTEPHALFLGIHGGKMLCSVK 309
QY 239 -----AEKLE-----AF-----DW----- 248
Db 310 GDETRILQLEAVNITDLSENKDKODKRFATIRSDSGPTTFSESAACPGWFLCTAMEADQPV 369
QY 249 -----LKAFY 253
Db 370 LTNMPDEGVWTKFY 384
RESULT 8
US-09-840-277-2
; Sequence 2, Application US/09840277
; Patent No. US20020158363A1
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: KOHNO, TADAHIKO
; APPLICANT: LACEY, DAVID LEE
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: INTEGRIN/ADHESION ANTAGONISTS
; FILE REFERENCE: A-688A
; CURRENT APPLICATION NUMBER: US/09/840,277
; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: 60/198,919
; PRIOR FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 60/201,394
; PRIOR FILING DATE: 2000-05-03
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-840-277-2
Query Match 86.6%; Score 1238; DB 9; Length 228;
Best Local Similarity 100.0%; Pred. No. 5.2e-32;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDKTHCTPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWY 60
Db 1 MDKTHCTPCPAPELLGGPSVFLFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWY 60
QY 61 DGVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKA 120
Db 61 DGVEVHNKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISKA 120
QY 121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 180
Db 121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 180
QY 181 SDGSFFLYSKLTVDKSRWQGNVFSCVMHEALHNHYTQKSLSLSPGK 228
Db 181 SDGSFFLYSKLTVDKSRWQGNVFSCVMHEALHNHYTQKSLSLSPGK 228
RESULT 9
US-09-847-249A-2
; Sequence 2, Application US/09847249A
; Publication No. US20030032588A1
; GENERAL INFORMATION:
; APPLICANT: MARSHALL, WILLIAM S.
; APPLICANT: STARK, KEVIN LEE
; TITLE OF INVENTION: GLUCAGON ANTAGONIST
; FILE REFERENCE: A-693
```

Query Match 86.6%; Score 1238; DB 9; Length 228;
Best Local Similarity 100.0%; Pred. No. 5.2e-32;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-09-847-249A-2

Sequence 2, Application US/09/847,249A
Publication No. US2003004070A1
APPLICANT: KOHNO, TADAHIKO
TITLE OF INVENTION: APO-AI/ALL PEPTIDE DERIVATIVES
CURRENT APPLICATION NUMBER: US/09/840,669B
CURRENT FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: 60/198,920
PRIOR FILING DATE: 2000-04-21
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 228
TYPE: PRT
ORGANISM: Homo sapiens

US-09-847-249A-2

Query Match 86.6%; Score 1238; DB 9; Length 228;
Best Local Similarity 100.0%; Pred. No. 5.2e-32;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-09-840-669B-2

Sequence 2, Application US/09/840,669B
Publication No. US2003004070A1
APPLICANT: KOHNO, TADAHIKO
TITLE OF INVENTION: APO-AI/ALL PEPTIDE DERIVATIVES
CURRENT APPLICATION NUMBER: US/09/840,669B
CURRENT FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: 60/198,920
PRIOR FILING DATE: 2000-04-21
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 228
TYPE: PRT
ORGANISM: Homo sapiens

Query Match 86.6%; Score 1238; DB 9; Length 228;
Best Local Similarity 100.0%; Pred. No. 5.2e-32;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-09-840-669B-2

Sequence 2, Application US/09/840,669B
Publication No. US2003004070A1
APPLICANT: KOHNO, TADAHIKO
TITLE OF INVENTION: APO-AI/ALL PEPTIDE DERIVATIVES
CURRENT APPLICATION NUMBER: US/09/840,669B
CURRENT FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: 60/198,920
PRIOR FILING DATE: 2000-04-21
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 228
TYPE: PRT
ORGANISM: Homo sapiens

Query Match 86.6%; Score 1238; DB 9; Length 228;
Best Local Similarity 100.0%; Pred. No. 5.2e-32;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-09-843-221A-2

Sequence 2, Application US/09/843,221A
Publication No. US20030039654A1
APPLICANT: KOSTENUK, PAUL
APPLICANT: LIU, CHUAN-FA
TITLE OF INVENTION: MODULATORS OF RECEPTORS FOR PARATHYROID HORMONE AND PARATHYROID HORMONE
CURRENT APPLICATION NUMBER: US/09/843,221A
CURRENT FILING DATE: 2001-04-26
PRIOR APPLICATION NUMBER: 60/266,673
PRIOR FILING DATE: 2001-02-06
PRIOR APPLICATION NUMBER: 60/214,860
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/200,053
PRIOR FILING DATE: 2000-04-27
NUMBER OF SEQ ID NOS: 170
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 228
TYPE: PRT
ORGANISM: Homo sapiens

Query Match 86.6%; Score 1238; DB 9; Length 228;
Best Local Similarity 100.0%; Pred. No. 5.2e-32;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-09-843-221A-2

Sequence 2, Application US/09/843,221A
Publication No. US20030039654A1
APPLICANT: KOSTENUK, PAUL
APPLICANT: LIU, CHUAN-FA
TITLE OF INVENTION: MODULATORS OF RECEPTORS FOR PARATHYROID HORMONE AND PARATHYROID HORMONE
CURRENT APPLICATION NUMBER: US/09/843,221A
CURRENT FILING DATE: 2001-04-26
PRIOR APPLICATION NUMBER: 60/266,673
PRIOR FILING DATE: 2001-02-06
PRIOR APPLICATION NUMBER: 60/214,860
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/200,053
PRIOR FILING DATE: 2000-04-27
NUMBER OF SEQ ID NOS: 170
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 228
TYPE: PRT
ORGANISM: Homo sapiens

Query Match 86.6%; Score 1238; DB 10; Length 228;
Best Local Similarity 100.0%; Pred. No. 5.2e-32;

US-09-847-712-2

Sequence 2, Application US/09/847,712
Patent No. US20020090646A1
GENERAL INFORMATION:
APPLICANT: MARSHALL, WILLIAM S.
APPLICANT: REYNOLDS, ANGELA
TITLE OF INVENTION: CALCITONIN-RELATED MOLECULES
FILE REFERENCE: A-684
CURRENT APPLICATION NUMBER: US/09/847,712
CURRENT FILING DATE: 2001-05-02
PRIOR APPLICATION NUMBER: 60/201,511
PRIOR FILING DATE: 2000-05-03
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 228
TYPE: PRT
ORGANISM: Homo sapiens

Query Match 86.6%; Score 1238; DB 10; Length 228;
Best Local Similarity 100.0%; Pred. No. 5.2e-32;

US-09-847-712-2

Sequence 2, Application US/09/847,712
Patent No. US20020090646A1
GENERAL INFORMATION:
APPLICANT: MARSHALL, WILLIAM S.
APPLICANT: REYNOLDS, ANGELA
TITLE OF INVENTION: CALCITONIN-RELATED MOLECULES
FILE REFERENCE: A-684
CURRENT APPLICATION NUMBER: US/09/847,712
CURRENT FILING DATE: 2001-05-02
PRIOR APPLICATION NUMBER: 60/201,511
PRIOR FILING DATE: 2000-05-03
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 228
TYPE: PRT
ORGANISM: Homo sapiens

US-10-215-297-2

```
Query Match      86.3%; Score 1234; DB 9; Length 229;
Best Local Similarity 99.6%; Pred. No. 7e-32; 0; Indels 0; Gaps 0;
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHTCPPCPAPPELLGGPSVFLFPKPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
Db :|||||
QY 2 VDKTHTCPPCPAPPELLGGPSVFLFPKPKDPTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 61
Db :|||||
QY 61 DGYEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
Db :|||||
QY 62 DGYEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 121
Db :|||||
QY 121 KGQPREPQVYTLPPSRDELTKNOVSLTCLVKGYFPSDIAVEWESNGQPENNYKTTTPVLD 180
Db :|||||
QY 181 SDGSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTQKSLSLSPCK 228
Db :|||||
QY 182 SDGSFFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTQKSLSLSPCK 229
Db :|||||
```

Search completed: April 21, 2003, 10:51:21
Job time : 14.1647 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:26 ; Search time 19.4118 seconds
(without alignments)
1307.428 Million cell updates/sec

Title: 2LINK7LINK7

Perfect score: 1430

Sequence: 1 MDKTHCTCPCPAPPELLGGPS.....AFDWLKAIFYDKVAEKLKEAF 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1233	86.2	255	4 S31866	Ig gamma-1 chain C
2	1233	86.2	330	1 GHU	Ig gamma-1 chain C
3	1227	85.8	374	2 S69339	Ig heavy chain V r
4	1180	82.5	234	2 PT0207	Ig gamma chain C r
5	1158	81.0	377	2 A29511	Ig gamma-3 chain C
6	1156	80.8	377	2 A60764	Ig gamma-3 chain C
7	1146	80.1	326	1 G2HU	Ig gamma-2 chain C
8	1142	79.9	327	1 G4HU	Ig gamma-4 chain C
9	1133	79.2	289	1 G3HUW1	Ig gamma-3 heavy c
10	921	64.4	323	1 GHRB	Ig gamma chain C r
11	917	64.1	328	2 I47160	Ig gamma 2b chain
12	917	64.1	328	2 I47159	Ig gamma 2a chain
13	911	63.7	277	2 I47162	Ig gamma 4 chain c
14	894	62.5	329	1 G2GP	Ig gamma-2 chain C
15	888	62.1	328	2 I47158	Ig gamma 1 chain c
16	885	61.9	328	2 I47161	Ig gamma 3 chain c
17	860	60.1	470	2 S22080	Ig heavy chain pre
18	858	60.0	472	2 S31459	Ig gamma-1 chain -
19	853	59.7	308	2 C30554	Ig heavy chain C r
20	848	59.3	329	1 G3MSC	Ig gamma-3 chain C
21	841	58.8	398	1 G3MSM	Ig gamma-3 chain C
22	840	58.7	333	2 PS0018	Ig gamma-2b chain
23	834	58.3	444	2 PC4436	monoclonal antibod
24	827	57.8	393	1 G1MSM	Ig gamma-1 chain C
25	825	57.7	326	2 PS0017	Ig gamma-1 chain C
26	823	57.6	324	1 G1MS	Ig gamma-1 chain C
27	819	57.3	399	1 G2MSAM	Ig gamma-1 chain
28	818	57.2	329	2 S00847	Ig gamma-2c chain
29	813	56.9	330	1 G2MSA	Ig gamma-2a chain

30	813	56.9	469	2 S37483	Ig gamma-2a chain
31	809	56.6	335	1 G2MSAB	Ig gamma-2a chain
32	799	55.9	446	2 S40295	Ig gamma-2a chain
33	793	55.5	322	2 PS0019	Ig gamma-2a chain
34	791	55.3	405	1 G2MSBM	Ig gamma-2b chain
35	787	55.0	474	1 G2MS11	Ig gamma-2b chain
36	774	54.1	327	2 S05611	Ig gamma-2 chain C
37	766	53.6	475	2 S01321	Ig gamma-2b chain
38	707	49.4	180	2 I46732	Ig gamma heavy cha
39	602	42.1	218	2 A36040	Ig heavy chain V-I
40	601	42.0	249	2 S69340	Ig heavy chain VHI
41	571	39.9	152	2 S14236	Ig gamma-1 chain C
42	493	34.5	26926	1 I38344	titin, cardiac mus
43	437	30.6	572	2 B45229	Ig y heavy chain (
44	422	29.5	548	2 S38864	Ig epsilon chain C
45	407	28.5	448	2 S03186	Ig heavy chain C r

ALIGNMENTS

RESULT 1

S31866

Ig gamma-1 chain C region - synthetic

C:Species: synthetic

A:Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli

C:Date: 06-Jan-1995 #sequence_revision 17-Mar-1997 #text_change 19-May-2000

C:Accession: S31866

R:Filpula, D.

submitted to the EMBL Data Library, February 1993

A:Description: Screening method for protein-protein interactions of cloned gene products

A:Reference number: S31866

A:Accession: S31866

A:Molecule type: mRNA

A:Residues: 1-255 <PIL>

A:Cross-references: EMBL:X70421; NID:g33068; FIDN:CAA49866.1; PID:g33069

C:Keywords: Immunoglobulin

F:1-22/Region: Escherichia coli outer membrane protein A precursor

F:23-255/Region: human Ig gamma-1 chain C region

Query Match

Best Local Similarity 86.2%; Score 1233; DB 4; Length 255;

Mismatches 0; Mismatches 0; Indels 0; Gaps 0;

Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DKTHCTCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVYVDSHEDPEVKFNMYVD 61

DB 29 DKTHCTCPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVYVDSHEDPEVKFNMYVD 88

QY 62 GVEVHNAKTPREEQYNSTYRVYVSVLTVLHODWLNCKEYKCKVSNKALPAPIEKTISKAK 121

DB 89 GVEVHNAKTPREEQYNSTYRVYVSVLTVLHODWLNCKEYKCKVSNKALPAPIEKTISKAK 148

QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 181

DB 149 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 208

QY 182 DGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNNHVTQKLSLSPGK 228

DB 209 DGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNNHVTQKLSLSPGK 255

RESULT 2

GHU

Ig gamma-1 chain C region - human

C:Species: Homo sapiens (man)

C:Date: 31-Jan-1981 #sequence_revision 18-Aug-1982 #text_change 16-Jul-1999

C:Accession: A93433; S36861; S3887; B90563; A90564; B91668; A91723; A02146

R:Ellison, J.W.; Berson, B.J.; Hood, L.E.

Nucleic Acids Res. 10, 4071-4079, 1982

A:Title: The nucleotide sequence of a human immunoglobulin C-gamma1 gene.

A:Reference number: A93433; MUID:82274238; PMID:6287432

A:Accession: A93433

A:Molecule type: DNA

A:Residues: 1-330 <ELL>
A:Cross-references: EMBL:Z17370
A:Note: this sequence has the Glm(17) allotypic marker, 97-Lys and the Glm(1) markers,
A:Note: Lys-330 is removed after translation
R:Harris, L.J.
submitted to the EMBL Data Library, October 1992
A:Reference number: S33904
A:Accession: S36861
A:Molecule type: DNA
A:Residues: 2-330 <HAR>
A:Cross-references: EMBL:Z17370
R:Takahashi, N.; Ueda, S.; Obata, M.; Nikaido, T.; Nakai, S.; Honjo, T.
Cell 29, 671-679, 1982
A:Title: Structure of human immunoglobulin gamma genes: Implications for evolution of a
A:Reference number: S33887; MUID:83001943; PMID:6811139
A:Accession: S33887
A:Molecule type: DNA
A:Residues: 88-113;235-330 <TAK>
A:Cross-references: EMBL:Z17370
R:Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman,
Biochemistry 9, 3161-3170, 1970
A:Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid sequen
A:Reference number: A90563; MUID:71064024; PMID:5489771
A:Contents: myeloma protein Eu
A:Accession: B90563
A:Molecule type: protein
A:Residues: 1-96,'R',98-135 <CUN>
A:Note: this sequence has the Glm(3) marker, 97-Arg
R:Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.
Biochemistry 9, 3171-3181, 1970
A:Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid sequen
A:Reference number: A90564; MUID:71064025; PMID:5530842
A:Contents: Eu
A:Accession: A90564
A:Molecule type: protein
A:Residues: 136-154,'Q',156-165,'Q',167-176,'Q',178-184,'N',196-197,'D',199-238,'E',240,
A:Note: this sequence has the Glm(non-1) markers, 239-Glu and 241-Met
R:Ponstingl, H.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976
A:Title: Die Primärstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein Nie).
igen Primärstruktur.
A:Reference number: A91668; MUID:77070269; PMID:826475
A:Contents: myeloma protein Nie
A:Accession: B91668
A:Molecule type: protein
A:Residues: 1-34,'Q',36-96,'K',98-115,'Q',117-197,'D',199-238,'D',240,'L',242-268,'E',27
A:Note: this sequence has the Glm(17) and Glm(1) markers
R:Schmidt, W.E.; Jung, H.D.; Palm, W.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 364, 713-747, 1983
A:Title: Die Primärstruktur des kristallisierten monoklonalen Immunglobulins IgG1 KOI
A:Reference number: A91723; MUID:83289131; PMID:6884994
A:Contents: myeloma protein KOI; disulfide bonds
A:Accession: A91723
A:Molecule type: protein
A:Residues: 1-96,'R',98-197,'D',199-238,'E',240,'M',242-266,'D',268-271,'D',273-330 <SCH
A:Note: this sequence has the Glm(3) and Glm(non-1) markers
R:Gall, W.E.; Edelman, G.M.
Biochemistry 9, 3188-3196, 1970
A:Title: The covalent structure of a human gammaG-immunoglobulin. X. Intrachain disulfid
A:Reference number: A90565; MUID:71064027; PMID:4923144
A:Contents: annotation; disulfide bonds
R:Dreker, L.; Schwarz, J.; Reichel, W.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 357, 1515-1540, 1976
A:Title: Rule of antibody structure. The primary structure of monoclonal IgG1 immunoglob
enbromide cleavage products, and the disulfide bridges.
A:Reference number: A91667; MUID:77070267; PMID:1002129
A:Contents: annotation; disulfide bonds
C:Genetics:
A:Gene: IGHG1
A:Cross-references: GDB:I20085; OMIM:147100
A:Map position: 14q32.33-14q32.33
A:Introns: 99/1; 114/1; 224/1
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap

chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:20-85/Domain: immunoglobulin homology <IM1>
F:137-206/Domain: immunoglobulin homology <IM2>
F:243-310/Domain: immunoglobulin homology <IM3>
F:27-83,144-204,250-308/Disulfide bonds: #status experimental
F:103/Disulfide bonds: interchain (to light chain) #status experimental
F:109,112/Disulfide bonds: interchain (to heavy chain) #status experimental
F:180/Binding site: carboxydrate (Asn) (covalent) #status experimental
Query Match 86.2%; Score 1233; DB 1; Length 330;
Best Local Similarity 100.0%; Pred. No. 6.5e-33;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 DKTHCPCPAPELLGGPSVFLPFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNMYVD 61
Db 104 DKTHCPCPAPELLGGPSVFLPFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNMYVD 163
Qy 62 GVEVHNAKTPREQYNSTYRVSVLTVLDHDLNGKEYCKVSNKALPAPIEKTISKAK 121
Db 164 GVEVHNAKTPREQYNSTYRVSVLTVLDHDLNGKEYCKVSNKALPAPIEKTISKAK 223
Qy 122 GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 181
Db 224 GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 283
Qy 182 DGSFFLSKLTVDKSRWQQGNVSCVMHEALHNNHYTKLSLSPGK 228
Db 284 DGSFFLSKLTVDKSRWQQGNVSCVMHEALHNNHYTKLSLSPGK 330
RESULT 3
S69339
Ig heavy chain V region precursor - human
C:Species: Homo sapiens (man)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Dec-2000
C:Accession: S69339; S72664
R:Khamilichi, A.A.; Aucouturier, P.; Preud'homme, J.L.; Cogne, M.
Eur. J. Biochem. 229, 54-60, 1995
A:Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease.
A:Reference number: S69339; MUID:95262687; PMID:7744049
A:Accession: S69339
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-374 <KHA>
A:Cross-references: EMBL:X81695
R:Khamilichi, A.A.
submitted to the EMBL Data Library, September 1994
A:Reference number: S72664
A:Accession: S72664
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-140,'C',142-374 <KHA>
A:Cross-references: EMBL:X81695
C:Superfamily: immunoglobulin C region; immunoglobulin homology
Query Match 85.8%; Score 1227; DB 2; Length 374;
Best Local Similarity 99.1%; Pred. No. 1.6e-32;
Matches 225; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 2 DKTHCPCPAPELLGGPSVFLPFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNMYVD 61
Db 148 DKTHCPCPAPELLGGPSVFLPFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNMYVD 207
Qy 62 GVEVHNAKTPREQYNSTYRVSVLTVLDHDLNGKEYCKVSNKALPAPIEKTISKAK 121
Db 208 GVEVHNAKTPREQYNSTYRVSVLTVLDHDLNGKEYCKVSNKALPAPIEKTISKAK 267
Qy 122 GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 181
Db 268 GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 327

A:Residues: 1-326 <ELL>
A:Cross-references: GB:V00554; GB:J00230; NID:g32759; PIDN:CAB58438.1; PID:g606056
A:Note: Lys-326 is probably removed posttranslationally
R:Wang, A.C.; Tung, E.; Fudenberg, H.H.
J. Immunol. 125, 1048-1054, 1980
A:Title: The primary structure of a human IgG2 heavy chain: genetic, evolutionary, and
A:Reference number: A92809; MUID:81007873; PMID:6774012
A:Contents: myeloma protein til
A:Accession: A92809
A:Molecule type: protein
A:Residues: 1-19, 'Q', 21-57, 'Z', 59, 'A', 61-193, 'D', 195-325 <NAN>
A:Note: Trp-156 is at or near the complement-binding site
R:Connell, G.E.; Parr, D.M.; Hofmann, T.
Can. J. Biochem. 57, 758-767, 1979
A:Title: The amino acid sequences of the three heavy chain constant region domains of a
A:Reference number: A90752; MUID:80001357; PMID:113060
A:Contents: myeloma protein zle
A:Accession: A90752
A:Molecule type: protein
A:Residues: 1-24, 'E', 26-57, 'EV', 60-85, 132-171, 'ZZZ', 175, 'B', 177-193, 'D', 195-196, 'Q', 198-
A:Note: This sequence has since been revised
R:Hofmann, T.; Parr, D.M.
Mol. Immunol. 16, 923-925, 1979
A:Title: A note on the amino acid sequence of residues 381-391 of human immunoglobulin G
A:Reference number: A93132; MUID:80114419; PMID:118920
A:Contents: zle
A:Accession: A93132
A:Molecule type: protein
A:Residues: 238-275 <HOF>
R:Hofmann, T.; Parr, D.M.
submitted to the Atlas, March 1980
A:Reference number: A94591
A:Contents: annotation; zle, revisions to residues 25, 59, 60, and 264-268
A:Note: The revised sequence differs from that shown in having 60-Ala and in the amidati
ned
R:Milstein, C.; Frangione, B.
Biochem. J. 121, 217-225, 1971
A:Title: Disulfide bridges of the heavy chain of human immunoglobulin G2.
A:Reference number: A90253; MUID:72033500; PMID:4940472
A:Contents: annotation; myeloma protein Sa, disulfide bonds
R:Frangione, B.; Milstein, C.; Pink, J.R.L.
Nature 221, 145-148, 1969
A:Title: Structural studies of immunoglobulin G.
A:Reference number: A93157; MUID:69064124; PMID:5782707
A:Contents: annotation; Sa, disulfide bonds
C:Genetics:
A:Gene: GDB:IGHG2
A:Cross-references: GDB:119338; OMIM:147110
A:Map position: 14q32.33-14q32.33
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:20-85/Domain: immunoglobulin homology <IM1>
F:133-202/Domain: immunoglobulin homology <IM2>
F:239-306/Domain: immunoglobulin homology <IM3>
F:14/Disulfide bonds: interchain (to light chain) #status experimental
F:27-83,140-200,246-304/Disulfide bonds: #status experimental
F:102,103,109/Disulfide bonds: interchain (to heavy chain) #status experimental
F:176/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 80.1%; Score 1146; DB 1; Length 326;
Best Local Similarity 90.6%; Pred. No. 4.6e-30;
Matches 212; Conservative 9; Mismatches 6; Indels 7; Gaps 2;
QY 1 MDKT-----HPCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVDSHEDPEV 54
DB 94 VDKTVERCCVCCPCPCAPP-VAGSVFLFPPKPKDTLMISRTPEVTCVVDVDSHEDPEV 152
QY 55 KFNWYVDGVGVHNAKTPREEQYNSTYRVSVLTVLDHQLNGKEYKCKVSNKALPAPIE 114
DB 153 QFNWYVDGVGVHNAKTPREEQNFSTYRVSVLTVLDHQLNGKEYKCKVSNKALPAPIE 212

QY 115 KTISKAKGQPREPVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKT 174
DB 213 KTISKAKGQPREPVYTLPPSRDEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKT 272
QY 175 TTPVLDSDGSEFFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPGK 228
DB 273 TTPMLDSDGSEFFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPGK 326
RESULT 8
GAMU
Ig gamma-4 chain C region - human
C:Species: Homo sapiens (man)
C:Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 16-Jul-1999
C:Accession: A90933; A90249; A02150
R:Ellison, J.; Buxbaum, J.; Hood, L.
DNA 1, 11-18, 1981
A:Title: Nucleotide sequence of a human immunoglobulin C-gamma4 gene.
A:Reference number: A90933; MUID:83157104; PMID:6299662
A:Accession: A90933
A:Molecule type: DNA
A:Residues: 1-327 <ELL>
A:Note: The sequence was determined from the germline gene
R:Pink, J.R.L.; Buttery, S.H.; De Vries, G.M.; Milstein, C.
Biochem. J. 117, 33-47, 1970
A:Title: Human immunoglobulin subclases. Partial amino acid sequence of the constan
A:Reference number: A90249; MUID:70207560; PMID:4192699
A:Accession: A90249
A:Molecule type: protein
A:Residues: 1-30; 81-326 <PIN>
C:Genetics:
A:Gene: GDB:IGHG4
A:Cross-references: GDB:119340; OMIM:147130
A:Map position: 14q32.33-14q32.33
A:Introns: 99/1; 111/1; 221/1
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (l
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into (l
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:20-85/Domain: immunoglobulin homology <IM1>
F:99-110/Region: hinge
F:134-203/Domain: immunoglobulin homology <IM2>
F:240-307/Domain: immunoglobulin homology <IM3>
F:14/Disulfide bonds: interchain (to light chain) #status experimental
F:27-83,141-201,247-305/Disulfide bonds: #status predicted
F:106,109/Disulfide bonds: interchain (to heavy chain) #status experimental
F:177/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 79.9%; Score 1142; DB 1; Length 327;
Best Local Similarity 84.7%; Pred. No. 6.3e-30;
Matches 211; Conservative 9; Mismatches 6; Indels 23; Gaps 1;
QY 3 KTHT-----CPCPAPELLGGPSVFLFPPKPKDTLMISRTPE 39
DB 79 KTYCNVDHKNPKNTKVRVESKYGPPCPCPAPELLGGPSVFLFPPKPKDTLMISRTPE 138
QY 40 VTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTPREEQYNSTYRVSVLTVLDHQLNGKE 99
DB 139 VTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTPREEQNFSTYRVSVLTVLDHQLNGKE 198
QY 100 YKCKVSNKALPAPIEKTISKAKGQPREPVYTLPPSRDELTKNQVSLTCLVKGFYPSDIA 159
DB 199 YKCKVSNKALPAPIEKTISKAKGQPREPVYTLPPSRDELTKNQVSLTCLVKGFYPSDIA 258
QY 160 VEWESNGQPENNYKTTPPVLDSDGSEFFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTQ 219
DB 259 VEWESNGQPENNYKTTPPVLDSDGSEFFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTQ 318
QY 220 KSLSLSPGK 228
DB 319 KSLSLSPGK 327


```
Db 151 FTWYINNEQVTRAPPLREQOQNSTIRVSTLPIHQDWLRGKFKCKVHNKALPAPIEK 210
QY 116 TISKAKQPREPOVYTLPPSRDELTKNOVSLTCLVKGFPSTIAVWESNGQPPENNYKTT 175
Db 211 TISKARQPLEPKVYVYNGPRELSRSVSLTCHNGFYPSPDISVWEKNGKAEADNYKT 270
QY 176 PVLDSGDFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTKSLSPGK 228
Db 271 PAVLDSGDFLYKLSVPTSEWQGDVFTCSVMHEALHNHYTKSLSPGK 323

RESULT 11
I47160
Ig gamma 2b chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47160
R:Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a
A:Reference number: I47158; MUID:95015845; PMID:7930579
A:Accession: I47160
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-328 <KAC>
A:Cross-references: EMBL:U03780; NID:g433125; PIDN:AAAS2218.1; PID:g433126
C:Genetics:
A:Gene: IgG2b
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 64.1%; Score 917; DB 2; Length 328;
Best Local Similarity 65.4%; Pred. No. 1.7e-22;
Matches 166; Conservative 32; Mismatches 24; Indels 32; Gaps 4;

QY 3 KTHTC-----PPCP-----APELLGSPVFLFPKPKDTLMISR 134
Db 79 KSYTCNVNHPATTTKVDKRVGKTKPPCPCIPACESP----GPSVFIFPPKPKDTLMISR 134
QY 37 TPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQNSTYRVVSVLTVLHQDWL 96
Db 135 TPQVTCVVVDVSOENPEVQFSWYVDGVEVHTAQTREKEQFNSTYRVVSVLPIQHODWLN 194
QY 97 GREYCKCKVSNKALPAPIETKISKAKQPREPOVYTLPPSRDELTKNOVSLTCLVKGFP 156
Db 195 GKEFKCKVNNKDLPAITRIISKAKQTPREPQVYTLPPHAEELSRKSVITCLVIGFY 254
QY 157 DIAVESNGQ--PENNYKTTTPVLDSGDFLYSKLTVDKSRWQOGNVFSCVMHEALH 214
Db 255 DIDVEQRNGQPEPEGNRYRTTPQDDVDGTYFLYSKFSVDKASWQGGGIFQCAVMHEALH 314
QY 215 NHYTKSLSPGK 228
Db 315 NHYTKSLSPGK 328

RESULT 13
I47162
Ig gamma 4 chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47162
R:Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of
A:Reference number: I47158; MUID:95015845; PMID:7930579
A:Accession: I47162
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-277 <KAC>
A:Cross-references: EMBL:U03782; NID:g433129; PIDN:AAAS2220.1; PID:g433130
C:Genetics:
A:Gene: IgG4
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:82-151/Domain: immunoglobulin homology <IMM>

Query Match 63.7%; Score 911; DB 2; Length 277;
Best Local Similarity 66.4%; Pred. No. 1.5e-22;
Matches 166; Conservative 31; Mismatches 29; Indels 24; Gaps 4;

QY 3 KTHTC-----PPCP-----APELLG--GPSVFLFPKPKDTLMISRTPEV 40
Db 28 KSYTCNVNHPATTTKVDKRVGKTKPPCPCIPACESPFIFFPKPKDTLMISRTPKV 87
QY 41 TCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQNSTYRVVSVLTVLHQDWLNGREY 100
Db 88 TCVVVDVSOENPEVQFSWYVDGVEVHTAQTREKEQFNSTYRVVSVLPIQHODWLN 147
QY 101 KCKVSNKALPAPIETKISKAKQPREPOVYTLPPSRDELTKNOVSLTCLVKGFP 160
Db 148 KCKVNNKDLPAITRIISKAKQTPREPQVYTLPPTEELSRKSVITCLVIGFY 207
QY 161 EWESNGQ--PENNYKTTTPVLDSGDFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYT 218
Db 208 EWQRNGQPEPEGNRYRTTPQDDVDGTYFLYSKLAVDKASWQGGTFFQCAVMHEALHNHYT 267
QY 219 QKSLSPGK 228
Db 268 QKSIFKTPGK 277
```

```
Db 151 FTWYINNEQVTRAPPLREQOQNSTIRVSTLPIHQDWLRGKFKCKVHNKALPAPIEK 210
QY 116 TISKAKQPREPOVYTLPPSRDELTKNOVSLTCLVKGFPSTIAVWESNGQPPENNYKTT 175
Db 211 TISKARQPLEPKVYVYNGPRELSRSVSLTCHNGFYPSPDISVWEKNGKAEADNYKT 270
QY 176 PVLDSGDFLYSKLTVDKSRWQOGNVFSCVMHEALHNHYTKSLSPGK 228
Db 271 PAVLDSGDFLYKLSVPTSEWQGDVFTCSVMHEALHNHYTKSLSPGK 323

RESULT 11
I47160
Ig gamma 2b chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47160
R:Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a
A:Reference number: I47158; MUID:95015845; PMID:7930579
A:Accession: I47160
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-328 <KAC>
A:Cross-references: EMBL:U03780; NID:g433125; PIDN:AAAS2218.1; PID:g433126
C:Genetics:
A:Gene: IgG2b
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 64.1%; Score 917; DB 2; Length 328;
Best Local Similarity 65.4%; Pred. No. 1.7e-22;
Matches 166; Conservative 32; Mismatches 24; Indels 32; Gaps 4;

QY 3 KTHTC-----PPCP-----APELLGSPVFLFPKPKDTLMISR 134
Db 79 KSYTCNVNHPATTTKVDKRVGKTKPPCPCIPACESP----GPSVFIFPPKPKDTLMISR 134
QY 37 TPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQNSTYRVVSVLTVLHQDWL 96
Db 135 TPQVTCVVVDVSOENPEVQFSWYVDGVEVHTAQTREKEQFNSTYRVVSVLPIQHODWLN 194
QY 97 GREYCKCKVSNKALPAPIETKISKAKQPREPOVYTLPPSRDELTKNOVSLTCLVKGFP 156
Db 195 GKEFKCKVNNKDLPAITRIISKAKQTPREPQVYTLPPHAEELSRKSVITCLVIGFY 254
QY 157 DIAVESNGQ--PENNYKTTTPVLDSGDFLYSKLTVDKSRWQOGNVFSCVMHEALH 214
Db 255 DIDVEQRNGQPEPEGNRYRTTPQDDVDGTYFLYSKFSVDKASWQGGGIFQCAVMHEALH 314
QY 215 NHYTKSLSPGK 228
Db 315 NHYTKSLSPGK 328

RESULT 12
I47159
Ig gamma 2a chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47159
R:Kacskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a
A:Reference number: I47158; MUID:95015845; PMID:7930579
A:Accession: I47159
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-328 <KAC>
A:Cross-references: EMBL:U03779; NID:g433123; PIDN:AAAS2217.1; PID:g433124
C:Genetics:
A:Gene: IgG2a
```


GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:21 ; Search time 10.0941 Seconds
(without alignments)
1084.766 Million cell updates/sec

Title: 2LINK7LINK7

Perfect score: 1430

Sequence: 1 MDKTHTCPPCPAPELLGPPS.....AFDWLKAFYDKVAEKLKEAF 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1233	86.2	330	1	GCL_HUMAN
2	1146	80.1	326	1	GC2_HUMAN
3	1142	79.9	327	1	GC4_HUMAN
4	1138	79.6	290	1	GC3_HUMAN
5	921	64.4	323	1	GC_RABIT
6	894	62.5	329	1	GC2_CAVPO
7	848	59.3	329	1	GC3_MOUSE
8	841	58.8	398	1	GC3M_MOUSE
9	840	58.7	333	1	GC3M_MOUSE
10	827	57.8	393	1	GC1M_MOUSE
11	825	57.7	326	1	GC1_RAT
12	823	57.6	324	1	GC1_MOUSE
13	819	57.3	399	1	GC4M_MOUSE
14	818	57.2	329	1	GC2_RAT
15	813	56.9	330	1	GC4A_MOUSE
16	809	56.6	335	1	GC4A_MOUSE
17	793	55.3	322	1	GC4_RAT
18	791	55.3	405	1	GC5M_MOUSE
19	785	54.9	336	1	GC2_MOUSE
20	387	26.7	421	1	EPC_MOUSE
21	382	26.7	454	1	MUC_HUMAN
22	372	26.0	429	1	EPC_RAT
23	370	25.9	391	1	MUCB_HUMAN
24	370	25.9	455	1	MUC_MOUSE
25	369	25.8	428	1	EPC_HUMAN
26	369	25.8	476	1	MUCM_MOUSE
27	367	25.7	458	1	MUC_RABIT
28	366	25.6	479	1	MUCM_RABIT
29	356	24.9	454	1	MUC_MESAU
30	349	24.4	450	1	MUC_CANFA
31	347	24.3	457	1	MUC_SUNMU
32	336	23.5	4969	1	RYR2_RABIT
33	336	23.5	6486	1	TYCC_BACBR

RESULT 1

ID	GCL_HUMAN	STANDARD	PRT	330 AA
AC	P01857			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Ig gamma-1 chain C region.			
GN	IGHG1			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
[1]				
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=82274238; PubMed=6287432;			
RA	Edlison J.W., Berson B.J., Hood L.E.;			
RT	"The nucleotide sequence of a human immunoglobulin C gamma1 gene."			
RL	Nucleic Acids Res. 10:4071-4079(1982).			
[2]				
RP	SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).			
RX	MEDLINE=71064024; PubMed=5489771;			
RA	Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,			
RA	Waxdal M.J., Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. VII. Amino			
RL	acid sequence of heavy-chain cyanogen bromide fragments H1-H4."			
RL	Biochemistry 9:3161-3170(1970).			
[3]				
RP	SEQUENCE OF 136-329 (EU).			
RX	MEDLINE=71064025; PubMed=5530842;			
RA	Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,			
RA	Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. 8. Amino			
RL	acid sequence of heavy-chain cyanogen bromide fragments H5-H7."			
RL	Biochemistry 9:3171-3181(1970).			
[4]				
RP	SEQUENCE (MYELOMA PROTEIN NIE).			
RX	MEDLINE=77070269; PubMed=826475;			
RA	Ponstingl H., Hilschmann N.;			
RT	"The rule of antibody structure. The primary structure of a			
RT	monoclonal IgG1 immunoglobulin (myeloma protein Nie). III. The			
RT	chymotryptic peptides of the H-chain, alignment of the tryptic			
RL	peptides and discussion of the complete structure."			
RL	Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).			
[5]				
RP	SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.			
RX	MEDLINE=83289131; PubMed=6884994;			
RA	Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;			
RT	"Three-dimensional structure determination of antibodies. Primary			
RT	structure of crystallized monoclonal immunoglobulin IgG1 KOL, I."			
RL	Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).			
[6]				
RP	DISULFIDE BONDS.			
RX	MEDLINE=71064027; PubMed=4923144;			
RA	Gall W.E., Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. X.			

Q92736 homo sapien
P11716 oryctolagus
P21817 homo sapien
P01875 gallus gall
P45443 neurospora
P23098 tripneustes
Q39565 chlamydomon
P23085 heterodontu
P20929 homo sapien
P04114 homo sapien
P39057 anthocidari
Q39575 chlamydomon

ALIGNMENTS

FT Intrachain disulfide bonds. ";
RL Biochemistry 9:3188-3196(1970).
RN [7]
RP DISULFIDE BONDS.
RX MEDLINE=77070267; PubMed=1002129;
RA Dreker L., Schwarz J., Reichel W., Hilschmann N.;
RA "Rule of antibody structure. The primary structure of a monoclonal
RT IgG1 immunoglobulin (myeloma protein Nie). I: Purification and
RT characterization of the protein, the L- and H-chains, the
RT cyanogen bromide cleavage products, and the disulfide bridges. ";
RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RX MEDLINE=81208100; PubMed=7236608;
RA Deisenhofer J.;
RA "Crystallographic refinement and atomic models of a human Fc fragment
RT aureus at 2.9- and 2.8-A resolution. ";
RT Biochemistry 20:2361-2370(1981).
RL -1- MISCELLANEOUS: NIE HAS THE GIM(17) ALLOTYPIC MARKER, 97-K, & THE
CC GIM(1) MARKERS, 239-D & 241-L. KOL & EU SEQUENCES HAVE THE GIM(3)
CC MARKER & THE GIM (NON-1) MARKERS.
CC -1- MISCELLANEOUS: NIE ALSO DIFFERS IN THE AMIDATION STATES OF
CC 35,116,198,269 & 272.
CC -1- MISCELLANEOUS: EU ALSO DIFFERS IN THE AMIDATION STATES OF RESIDUES
CC 155, 166, 177, 195, 198, 269, AND 272 AND IN THE ORDER OF RESIDUES
CC 268-272.
CC -1- MISCELLANEOUS: KOL ALSO DIFFERS IN THE AMIDATION STATES OF
CC RESIDUES 198,267&272.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J00228; AAC82527.1; ALT_INIT.
DR PIR; A02146; GHU.
DR PDB; 1FC1; 15-JUL-92.
DR PDB; 1FC2; 15-JUL-92.
DR Genem; HGNC:5525; IGHL1.
DR MIN; 147100; -.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003600; Ig-like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; Ig-like; 1.
DR SMART; SM00407; IGc1; 2.
DR PROSITE; PS00290; IG_MHC; 2.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW 3D-structure.
FT NON_TER 1 1
FT DOMAIN 1 98 CH1.
FT DOMAIN 99 110 HINGE.
FT DOMAIN 111 223 CH2.
FT DOMAIN 224 330 CH3.
FT DISULFID 27 83
FT DISULFID 103 103 INTERCHAIN (WITH LIGHT CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH HEAVY CHAIN).
FT DISULFID 144 204
FT DISULFID 250 308
FT CARBOHYD 180 180
FT MOD_RES 330 330
FT VARIANT 97 97
FT N-LINKED (GLNAC...).
FT REMOVED POST-TRANSLATIONALLY.
FT K -> R (IN GIM(3) MARKER).
FT /FTIG-VAR_003886.
FT D -> E (IN GIM(NON-1) MARKER).
FT /FTIG-VAR_003887.
FT L -> M (IN GIM(NON-1) MARKER).
FT /FTIG-VAR_003888.
FT STRAND 123 126

FT HELIX 130 134
FT TURN 136 137
FT STRAND 141 148
FT TURN 158 162
FT STRAND 163 164
FT TURN 165 166
FT STRAND 175 178
FT STRAND 183 190
FT HELIX 193 197
FT TURN 198 199
FT STRAND 202 206
FT STRAND 215 219
FT STRAND 227 227
FT STRAND 230 234
FT HELIX 238 240
FT TURN 241 242
FT STRAND 245 256
FT STRAND 260 266
FT TURN 267 268
FT STRAND 269 270
FT STRAND 274 276
FT STRAND 280 281
FT TURN 283 284
FT STRAND 287 296
FT HELIX 297 301
FT TURN 302 303
FT STRAND 306 312
FT TURN 313 314
FT TURN 316 317
FT STRAND 320 324
SQ SEQUENCE 330 AA; 36106 MW; 3770EE106C2FA33D CRC64;
Query Match 86.2%; Score 1233; DB 1; Length 330;
Best Local Similarity 100.0%; Pred. No. 1.le-33;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 DKTHTCPCPAPELLGGPSVFLFPPPKPKDTLMISRTPEVTCVVYVDSHEDPEVKFNWYVD 61
DB 104 DKTHTCPCPAPELLGGPSVFLFPPPKPKDTLMISRTPEVTCVVYVDSHEDPEVKFNWYVD 163
QY 62 GVEVHNAKTPREEQYNSTYRVYSVLTVLDHQLNGKEKCKVSNKALPAPIEKTIISKAK 121
DB 164 GVEVHNAKTPREEQYNSTYRVYSVLTVLDHQLNGKEKCKVSNKALPAPIEKTIISKAK 223
QY 122 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 181
DB 224 GQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 283
QY 182 DGSFFLYSKLTVDKSRWQQGNVSCFVSVMHEALHNHYTQKSLSLSPGK 228
DB 284 DGSFFLYSKLTVDKSRWQQGNVSCFVSVMHEALHNHYTQKSLSLSPGK 330.
RESULT 2
GC2_HUMAN
ID GC2_HUMAN STANDARD; PRT; 326 AA.
AC P01859;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
GN IGHG2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE OF 2-326 FROM N.A.
RX MEDLINE=82197621; PubMed=6804948;
RA Ellison J.W., Hood L.E.;
RT "Linkage and sequence homology of two human immunoglobulin gamma
RT heavy chain constant region genes. ";
RL Proc. Natl. Acad. Sci. U.S.A. 79:1984-1988(1982).

DR	Genew; HGNC:5526; IGHC2.
DR	MIM; 147110; .
DR	InterPro; IPR003006; Ig_MHC.
DR	InterPro; IPR003597; Ig-cl.
DR	InterPro; IPR003600; Ig_like.
DR	Pfam; PF00047; ig; 3.
DR	SMART; SM00410; Ig_like; 1.
DR	SMART; SM00407; IGcl; 2.
DR	PROSITE; PS00290; IG_MHC; 2.
KW	Immunoglobulin domain; Immunoglobulin C region.
FT	NON_TER 1 98
FT	DOMAIN 1 98
FT	DOMAIN 99 110
FT	DOMAIN 111 219
FT	DOMAIN 220 326
FT	FT DISULFID 14 14
FT	FT DISULFID 27 83
FT	FT DISULFID 102 102
FT	FT DISULFID 103 103
FT	FT DISULFID 106 106
FT	FT DISULFID 109 109
FT	FT DISULFID 140 200
FT	FT DISULFID 246 304
FT	FT SITE 156 156
FT	MOD_RES 326 326
FT	VARIANT 60 60
FT	CONFLICT 109 109
FT	SEQUENCE 326 AA; 35884 MW; 8310678C687CF9C CRC64;
QY	1 MDKT-----HTCCPCAPPELLGGPSVLEPPKPKDLMISRPEVTCVVVDVSHEDPEV 54
Db	94 VDKTVERKCCVECPPEAPP-VAGSFVLPFPKPKDLMISRPEVTCVVVDVSHEDPEV 154
QY	55 KFNMYVDGVEVHNNAKTPREQYNASTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIE 114
Db	153 QFNMYVDGVEVHNNAKTPREQFNASTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIE 211
QY	115 KTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKT 171
Db	213 KTISKTKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKT 273
QY	175 TTPVLDSDGSFFLYSKLTVDKSRWQQGNVSCSYMHHEALHNHYTQKSLSLSPGK 228
Db	273 TTPVLDSDGSFFLYSKLTVDKSRWQQGNVSCSYMHHEALHNHYTQKSLSLSPGK 326
RESULT 3	
GC4_HUMAN	STANDARD; PRT; 327 AA.
ID GC4_HUMAN	
AC P01861;	
DT 21-JUL-1986 (Rel. 01, Created)	
DT 21-JUL-1986 (Rel. 01, Last sequence update)	
DT 16-OCT-2001 (Rel. 40, Last annotation update)	
DE ig gamma-4 chain C region.	
GN IGHG4.	
OS Homo sapiens (Human).	
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	
OX NCBI_TaxID=9606;	
RP [1]	
RP SEQUENCE FROM N.A.	
RX MEDLINE=83157104; PubMed=6299662;	
RA Ellison J.W., Buxbaum J.N., Hood L.E.;	
RT "Nucleotide sequence of a human immunoglobulin C gamma 4 gene.";	
RL DNA 1:11-18(1981).	
RP [2]	
RP SEQUENCE OF 1-30 AND 81-326.	
RP MEDLINE=70207560; PubMed=4192699;	

RA Pink J.R.L., Buttery S.H., de Vries G.M., Milstein C.;
 RT "Human immunoglobulin subclasses. Partial amino acid sequence of the
 RL constant region of a gamma 4 chain.";
 CC Biochem. J. 117:33-47(1970).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: K01316; AAB5394.1; ALT_INIT.
 DR PIR: A02150; G4HU.
 DR HSP: P01842; 7FAB.
 DR Genew; HGNC:5528; IGHG4.
 DR MIM: 147130; -
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003597; Ig_c1.
 DR InterPro; IPR003600; Ig_like.
 DR Pfam: PF00047; Ig; 3.
 DR SMART; SM00410; Ig_like; 1.
 DR SMART; SM00407; IGc1; 2.
 DR PROSITE; PS00290; IG_MHC; 2.
 DR Immunoglobulin domain; Immunoglobulin C region.
 KW NON_TER 1
 FT DOMAIN 1 98
 FT CH1.
 FT HINGE.
 FT DOMAIN 99 110
 FT DOMAIN 111 220
 FT DOMAIN 221 327
 FT CH2.
 FT DISULFID 14 14
 FT DISULFID 27 83
 FT DISULFID 106 106
 FT DISULFID 109 109
 FT DISULFID 141 201
 FT DISULFID 247 305
 SQ SEQUENCE 327 AA; 35940 MW; 3EDBD811EF208E7A CRC64;
 Query Match 79.9%; Score 1142; DB 1; Length 327;
 Best Local Similarity 84.7%; Pred. No. 1.1e-30;
 Matches 211; Conservative 9; Mismatches 6; Indels 23; Gaps 1;
 QY 3 KTHT-----CPPCPAPELLGGPSVFLPPPKPKDTLMISRTPE 39
 Db 79 KTYTCNVDPKNTKVDKRVESKYGPCPCAPFLGGPSVFLPPPKPKDTLMISRTPE 138
 QY 40 VTCVVVDVSHEDPEVKFNKYDQVEVHNAKTKPREEQYNSTYRVYSLVTLVHQDLNGKE 99
 Db 139 VTCVVVDVSHEDPEVKFNKYDQVEVHNAKTKPREEQYNSTYRVYSLVTLVHQDLNGKE 198
 QY 100 YKCKVSNKALPAPIEKTISKAKGQPREPVYTLPPSRDELTKNQVSLTCLVKGFYPSDIA 159
 Db 199 YKCKVSNKGLPSSIEKTISKAKGQPREPVYTLPPSRDEMTKNQVSLTCLVKGFYPSDIA 258
 QY 160 VEWESNGPENNKTTPPVLDSDGFFLYSKLTVYDKSRWQQGNVFSCSVMHEALHNHYTQ 219
 Db 259 VEWESNGPENNKTTPPVLDSDGFFLYSKLTVYDKSRWQQGNVFSCSVMHEALHNHYTQ 318
 QY 220 KSLSLSPGK 228
 Db 319 KSLSLSLGK 327
 RESULT 4
 GC3_HUMAN
 ID GC3_HUMAN STANDARD; PRT; 290 AA.
 AC P01860;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ig gamma-3 chain C region (Heavy chain disease protein) (HDC).
 GN IGHG3.

OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE (DISEASE PROTEIN WIS).
 RX MEDLINE=81021548; PubMed=6774747;
 RA Frangione B., Rosenwasser E., Pirelli F., Franklin E.C.;
 RT "Primary structure of human gamma 3 immunoglobulin deletion mutant:
 RL gamma 3 heavy-chain disease protein wis.";
 RL Biochemistry 19:4304-4308(1980).
 RN [2]
 RP REVISIONS TO 12-97 (PROTEIN WIS).
 RX MEDLINE=77118561; PubMed=402363;
 RA Michaelson T.E., Frangione B., Franklin E.C.;
 RT "Primary structure of the 'hinge' region of human IgG3. Probable
 RL quadruplication of a 15-amino acid residue basic unit.";
 RL J. Biol. Chem. 252:883-889(1977).
 RN [3]
 RP REVISIONS TO 59-289 (PROTEIN WIS/DISEASE PROTEIN ZUC).
 RX MEDLINE=77021516; PubMed=823945;
 RA Wolfenstein-Todel C., Frangione B., Pirelli F., Franklin E.C.;
 RT "The amino acid sequence of 'heavy chain disease' protein ZUC.
 RL Structure of the FC fragment of immunoglobulin G3.";
 RL Biochem. Biophys. Res. Commun. 71:907-914(1976).
 RN [4]
 RP SEQUENCE FROM N.A. (DISEASE PROTEIN OMM).
 RX MEDLINE=82247835; PubMed=6808505;
 RA Alexander A., Steinmetz M., Barritault D., Frangione B.,
 RA Franklin E.C., Hood L., Buxbaum J.N.;
 RT "Gamma Heavy chain disease in man: cDNA sequence supports partial
 RL gene deletion model.";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:3260-3264(1982).
 CC -1- SUBUNIT: DIMER LINKED BY 12 DISULFIDE BONDS; IT HAS AN EXTRA
 CC INTERCHAIN DISULFIDE BOND AT POSITION 7 IN ADDITION TO THE 11
 CC NORMALLY PRESENT IN THE HINGE REGION.
 CC -1- MISCELLANEOUS: THE HEAVY CHAIN DISEASE PROTEIN WIS IS SHOWN.
 CC -1- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 42-76 WAS TAKEN FROM THE
 CC REF.2.
 CC -1- MISCELLANEOUS: DISEASE PROTEIN WIS IS LACKING MOST OF THE V REGION
 CC AND ALL OF THE CH1 REGION.
 CC -1- MISCELLANEOUS: DISEASE PROTEIN ZUC LACK MOST OF THE V REGION, ALL
 CC OF THE CH1 REGION, AND PART OF THE HINGE COMPARED WITH NORMAL
 CC GAMMA-3 HEAVY CHAINS.
 CC -1- MISCELLANEOUS: DISEASE PROTEIN OMM MAY REPRESENT AN ALLELIC FORM
 CC OR ANOTHER GAMMA CHAIN SUBCLASS.
 CC -1- MISCELLANEOUS: THE HINGE REGION IN GAMMA-3 CHAINS IS ABOUT FOUR
 CC TIMES AS LONG AS IN OTHER GAMMA CHAINS AND CONTAINS THREE
 CC IDENTICAL 15-RESIDUE SEGMENTS PRECEDED BY A SIMILAR 17-RESIDUE
 CC SEGMENT (12-28).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J00231; AAB52805.1; ALT_SEQ.
 DR PIR: A02149; G3HUWI.
 DR HSP: P01857; 1FC1.
 DR Genew; HGNC:5527; IGHG3.
 DR MIM: 147120; -
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003597; Ig_c1.
 DR InterPro; IPR003600; Ig_like.
 DR Pfam: PF00047; Ig; 2.
 DR SMART; SM00410; Ig_like; 1.
 DR SMART; SM00407; IGc1; 1.
 DR PROSITE; PS00290; IG_MHC; 1.
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein; Repeat.
 GN DOMAIN 12 73
 FT HINGE.

```

FT DOMAIN 74 183 CH2.
FT DOMAIN 184 289 CH3.
FT REPEAT 29 43
FT REPEAT 44 58
FT REPEAT 59 73
FT MOD_RES 1 1
FT CARBOHYD 6 6
FT DISULFID 7 7
FT DISULFID 24 24
FT DISULFID 27 27
FT DISULFID 33 33
FT DISULFID 39 39
FT DISULFID 42 42
FT DISULFID 48 48
FT DISULFID 54 54
FT DISULFID 57 57
FT DISULFID 63 63
FT DISULFID 69 69
FT DISULFID 72 72
FT CARBOHYD 140 140
FT MOD_RES 290 290
FT VARIANT 126 127
FT VARIANT 134 134
FT VARIANT 139 139
FT VARIANT 182 182
FT VARIANT 227 227
FT VARIANT 227 227
FT VARIANT 279 279
SQ SEQUENCE 290 AA; 32331 MW; E69CBC95705B2F46 CRC64;

Query Match 79.6%; Score 1138; DB 1; Length 290;
Best Local Similarity 76.5%; Pred. No. 1.1e-30;
Matches 208; Conservative 13; Mismatches 6; Indels 45; Gaps 1;

QY 2 DKTHT-----CPCPAPELL 16
Db 19 DTHTPCRCPEKSCDTPPCRCPEKSCDTPPCRCPEKSCDTPPCRCPAPELL 78
QY 17 GGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKPREQ 76
Db 79 GGPSVFLFPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNATKPREQ 138
QY 77 YNSTYRVSVTLVHQLDNLGKYEKCKVSNKALPAPIETKISKAKGQPREPPQVYITLPPSR 136
Db 139 FNSFTRVSVTLVHQLDNLGKYEKCKVSNKALPAPIETKISKAKGQPREPPQVYITLPPSR 198
QY 137 DEUTKNOVSLTCLVKGYFIPSDIAVEWESNGQPENNYKTPPPVLDSDGSFFLYSKLTVDKS 196
Db 199 EEMTKNOVSLTCLVKGYFIPSDIAVEWESNGQPENNYKTPPPVLDSDGSFFLYSKLTVDKS 258
QY 197 RWQGNVFCVSMHEALHNYHTOKSLSPGK 228
Db 259 RWQGNVFCVSMHEALHNYHTOKSLSPGK 290

RESULT 5
GC_RABIT STANDARD; PRT; 323 AA.
AC P01870;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma chain C region.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

```

```

OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84030930; PubMed=6313520;
RA Bernstein K.E., Alexander C.B., Mage R.G.;
RT "Nucleotide sequence of a rabbit IgG heavy chain from the recombinant
RL F-1 haplotype.";
RN Immunogenetics 18:387-397(1983).
[2]
RP SEQUENCE OF 1-128
RX MEDLINE=76135469; PubMed=1243651;
RA Pratt D.M., Mole L.E.;
RT "Sequence studies on the constant region of the Fd sections of rabbit
RL immunoglobulin G of different allotype.";
RN Biochem. J. 151:337-349(1975).
[3]
RP SEQUENCE OF 88-266 FROM N.A.
RX MEDLINE=83299917; PubMed=6193512;
RA Martens C.L., Moore K.W., Steinmetz M., Hood L., Knight K.L.;
RT "Heavy chain genes of rabbit IgG: isolation of a cDNA encoding gamma
RL heavy chain and identification of two genomic C gamma genes.";
RN Proc. Natl. Acad. Sci. U.S.A. 79:6018-6022(1982).
[4]
RP SEQUENCE OF 132-161.
RX MEDLINE=70110015; PubMed=5461106;
RA Fruchter R.G., Jackson S.A., Mole L.E., Porter R.R.;
RT "Sequence studies of the Fd section of the heavy chain of rabbit
RL immunoglobulin G.";
RN Biochem. J. 116:249-259(1970).
[5]
RP SEQUENCE OF 129-131 AND 155-322.
RA Hill R.L., Lebovitz H.E., Fellows R.E. Jr., Delaney R.;
RL (In) Killander J. (eds.);
RL Gamma globulins, Nobel symp. 3, pp.109-127, Almqvist and Wiksell,
RL Stockholm (1967).
CC -!- MISCELLANEOUS: REF.1 SEQUENCE HAS THE D12 ALLOTYPIC MARKER,
CC 104-THR, AND THE E14 MARKER, 185-THR. REF.3 HAS THE D11 AND E15
CC MARKERS AND REF.5 THE E15 MARKER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M16426; AAA31289.1; -.
DR PIR; A02161; GHRB.
DR HSSP; P01857; 1FC1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00407; Igc1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
DR Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT VARIANT 104 104 T -> M (IN D11 MARKER).
FT VARIANT 185 185 T -> A (IN E15 MARKER).
FT CONFLICT 48 48 N -> E (IN REF. 2).
FT CONFLICT 71 71 V -> VPV (IN REF. 2).
FT CONFLICT 144 144 Q -> E (IN REF. 3 AND 4).
FT CONFLICT 173 173 N -> D (IN REF. 5).
FT CONFLICT 187 187 Q -> E (IN REF. 3 AND 5).
FT CONFLICT 201 201 N -> D (IN REF. 5).
FT CONFLICT 218 218 Q -> E (IN REF. 5).
FT CONFLICT 233 233 N -> Q (IN REF. 5).
FT CONFLICT 246 246 E -> G (IN REF. 5).
FT CONFLICT 256 256 N -> D (IN REF. 5).
FT CONFLICT 260 260 N -> D (IN REF. 5).
FT CONFLICT 266 266 N -> D (IN REF. 5).
FT CONFLICT 280 280 Y -> W (IN REF. 5).
FT CONFLICT 284 284 N -> S (IN REF. 5).

```

SQ SEQUENCE 323 AA; 35404 MW; 69E8AA118D579A8B CRC64;
 Query Match 64.4%; Score 921; DB 1; Length 323;
 Best Local Similarity 71.7%; Pred. No. 2.5e-23;
 Matches 167; Conservative 29; Mismatches 32; Indels 5; Gaps 2;
 QY 1 MDKT---HTC--PPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 55
 DB 91 VDKTVPSTCTCKTCTCPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVK 150
 QY 56 FNNYVDGVEVHNKATPREQYVNSTRVVSVLVTLVHODWLNKGYCKVSKNKPALPIEK 115
 DB 151 FTWIIINNEQVTRAPPLRQEQNSTIRVSVTLPIHODWLNKGYCKVSKNKPALPIEK 210
 QY 116 TISKAGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVESNGQPPENYKTT 175
 DB 211 TISKAGQPLEPKVYVYTGPPRELSRSVSLTCLVINGFYPDSISVSEWKNKRAEDNYKTT 270
 QY 176 PVLVSDGSGFFLYSKLTVDKSRQOQNVFSCSYMHEALHNHYTKLSLSLSPGK 228
 DB 271 PAVLSDGSGFLYKLNKLSVPTSEWQGDVFTCSVMHEALHNHYTKLSISRSPGK 323
 RESULT 6
 GC2_CAVPO STANDARD; PRT; 329 AA.
 ID GC2_CAVPO
 AC P01862;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig gamma-2 chain C region.
 OC Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE OF 1-3.
 RA Trischmann T.M.;
 RL Submitted (APR-1975) to the PIR data bank.
 RN [2]
 RP SEQUENCE OF 4-68.
 RX MEDLINE=71058471; PubMed=5538606;
 RA Birshstein B.K., Hussain Q.Z., Cebra J.J.;
 RT "Structure of heavy chain from strain 13 guinea pig
 immunoglobulin-G(2). 3. Amino acid sequence of the region around the
 half-cysteine joining heavy and light chains.";
 RL Biochemistry 10:18-25(1971).
 RN [3]
 RP SEQUENCE OF 69-133 AND 312-329.
 RX MEDLINE=71058486; PubMed=5538616;
 RA Turner K.J., Cebra J.J.;
 RT "Structure of heavy chain from strain 13 guinea pig
 immunoglobulin-G(2). II. Amino acid sequence of the carboxyl-terminal
 and hinge region cyanogen bromide fragments.";
 RL Biochemistry 10:9-17(1971).
 RN [4]
 RP SEQUENCE OF 134-226.
 RX MEDLINE=75036072; PubMed=4429665;
 RA Tracey D.E., Cebra J.J.;
 RT "Primary structure of the CH2 homology region from guinea pig IgG2
 antibodies.";
 RL Biochemistry 13:4796-4803(1974).
 RN [5]
 RP SEQUENCE OF 227-311.
 RX MEDLINE=75036073; PubMed=4609467;
 RA Trischmann T.M., Cebra J.J.;
 RT "Primary structure of the CH3 homology region from guinea pig IgG2
 antibodies.";
 RL Biochemistry 13:4804-4811(1974).
 RN [6]
 RP DISULFIDE BONDS.
 RX MEDLINE=71058474; PubMed=4922544;
 RA Oliveira B., Lamm M.E.;

RT "Interchain disulfide bridges of guinea pig gamma-2-immunoglobulin.";
 RL Biochemistry 10:26-31(1971).
 CC !- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM POOLED SERUM OF STRAIN
 CC 13 INBRED GUINEA PIGS.
 DR PIR; A02151; G2GP.
 DR HSSP; P01842; 7FAB.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003597; Ig_c1.
 DR InterPro; IPR003600; Ig_like.
 DR Pfam; PF00047; Ig; 2.
 DR SMART; SM00410; Ig_like; 1.
 DR SMART; SM00407; IGc1; 2.
 DR PROSITE; PS00290; IG_MHC; 1.
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
 FT NON_TER 1 1
 FT DISULFID 16 16
 FT DISULFID 28 79
 FT DISULFID 105 105
 FT DISULFID 107 107
 FT DISULFID 110 110
 FT DISULFID 142 202
 FT CARBOHID 178 178
 FT DISULFID 248 308
 SQ SEQUENCE 329 AA; 36074 MW; 5D231B7164D1FBA9 CRC64;
 Query Match 62.5%; Score 894; DB 1; Length 329;
 Best Local Similarity 68.8%; Pred. No. 2.1e-22;
 Matches 165; Conservative 25; Mismatches 37; Indels 13; Gaps 2;
 QY 1 MDKT-----HTCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSH 49
 DB 90 VDKTVEPIRTZPBPCCTCPKPPENLGGPSVFLFPPKPKDTLMISLTPTVTCVVVDVSH 149
 QY 50 EPEVKFNWYVDCVEVHNKATPREQYVNSTRVVSVLVTLVHODWLNKGYCKVSKNKA 109
 DB 150 DEPEVQFTWFDNKPVGNAETPRVEQYNTFRVESVLPVHODWLNKGYCKVSKNKA 209
 QY 110 PAPIEKTISKAGQPREPOVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVESNGQP- 168
 DB 210 PAPIEKTISKAGAPMPDYTLPPSRDELTKNSKSVTCLINFFPADTHVEASNRVPV 269
 QY 169 -ENNYKTPPVLDSDGSGFFLYSKLTVDKSRQOQNVFSCSYMHEALHNHYTKLSLS 327
 DB 270 SEKEYKNTPIEDADGSGFLYKLVKSAWQGTVTCSVMHEALHNHYTKAISRSPG 329
 RESULT 7
 GC3_MOUSE STANDARD; PRT; 329 AA.
 ID GC3_MOUSE
 AC P22436;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ig gamma-3 chain C region, secreted form.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85027161; PubMed=6092053;
 RA Wells J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
 RA Tucker P.W., Blattner F.R.;
 RT "Structural analysis of the murine IgG3 constant region gene.";
 RL EMBO J. 3:2041-2046(1984).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```

CC EMBL; J00451; NOT_ANNOTATED_CDS.
DR PIR; B02156; G3MSC.
DR HSSP; P01857; 1FC1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; Ig_like; 1.
DR SMART; SM00407; IGcl; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
FT NON_TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
FT TRANSMEM 345 362 POTENTIAL.
FT DOMAIN 363 398 CYTOPLASMIC (POTENTIAL).
FT CONFLICT 333 342 E -> G (IN REF. 2).
FT CONFLICT 342 342 E -> Q (IN REF. 2).
FT CONFLICT 388 388 P -> F (IN REF. 2).
SQ SEQUENCE 329 AA; 36228 MW; F45827174182BAD6 CRC64;

Query Match 59.3%; Score 848; DB 1; Length 329;
Best Local Similarity 60.1%; Pred. No. 7.3e-21;
Matches 152; Conservative 35; Mismatches 39; Indels 27; Gaps 2;

QY 3 KTHTC-----PP---CPAPELLGGPSVFLFPPKPKDTLMIS 35
Db 77 QTVICNVHPASKTELKRIEPRIPKPTPPGSSCPGNILGGPSVFIFFPKPKDALMIS 136
QY 36 RTEVTCVVVDVSHEDPEVKFNVDGVEVHNAKTKPRREQYNSTYRVVSVLTVLHQDWL 95
Db 137 LTPKVTCTVVVDSDDDPDVHVSFVFNKVEVHTAQTQPREAQYNSTFRVVSALPIQHDWM 196
QY 96 NGREYKCKVSKALPAPIETKISKAGOPREPOVYITLPPSRDELTKNOVSLTCLVKGFP 155
Db 197 RGFEKCKVKNKALPAPIETKISKGRAGTQVYITPPPREQSKKVSITCLVTNFFS 256
QY 156 SDIAVESNGQPNKYKTPPVLDSGSEFLYSLKTVDKSRWQQGNVFCSCVMHEALHN 215
Db 257 EATSVEMERGEQDYKNTPPILDSGTYFLYSLKTVDTDSWLGQEIFTCVSVVHEALHN 316
QY 216 HYTKQSLSLSPGK 228
Db 317 HHTQKNLSRSPGK 329

RESULT 8
GC3M_MOUSE STANDARD; PRT; 398 AA.
AC P03987;
DT 23-OCT-1986 (Rel. 02, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-3 chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;
RA Weis J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBO J. 3:2041-2046(1984).
RN [2]
RP SEQUENCE OF 328-398 FROM N.A.
RX MEDLINE=84041483; PubMed=6314258;
RA Komaromy M., Clayton L., Rogers J., Robertson S., Kettman J.,
RA Wall R.;
RT "The structure of the mouse immunoglobulin in gamma 3 membrane gene
segment.";
RL Nucleic Acids Res. 11:6775-6785(1983).

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00451; AAB59655.1; -.
DR EMBL; V01526; CAA24767.1; ALT_SEQ.
DR PIR; A02155; G3MSC.
DR HSSP; P01857; 1FC1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; Ig_like; 1.
DR SMART; SM00407; IGcl; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
FT NON_TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
FT TRANSMEM 345 362 POTENTIAL.
FT DOMAIN 363 398 CYTOPLASMIC (POTENTIAL).
FT CONFLICT 333 342 E -> G (IN REF. 2).
FT CONFLICT 342 342 E -> Q (IN REF. 2).
FT CONFLICT 388 388 P -> F (IN REF. 2).
SQ SEQUENCE 398 AA; 43929 MW; CF7E264B50A41B95 CRC64;

Query Match 58.8%; Score 841; DB 1; Length 398;
Best Local Similarity 55.0%; Pred. No. 2.3e-20;
Matches 153; Conservative 37; Mismatches 44; Indels 44; Gaps 3;

QY 8 PP---CPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNKYVDGVE 64
Db 106 PGSSCPPGNILGGPSVFIFFPKPKDALMISLTPKVTCTVVVDSEDDPDVHVSFVFNKE 165
QY 65 VHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIETKISKAGQP 124
Db 166 VHTAQTQPREAQYNSTFRVVSALPIQHDWMRGKEFKVNNKALPAPIETISKPKGRA 225
QY 125 REPOVYITLPPSRDELTKNOVSLTCLVKGFPSPDIATVWESNGOPENNYKTPPVLDSGS 184
Db 226 QTPQVYITPPPREQSKKVSITCLVTNFFSEALISWEERGELEQDYKNTPTPILDSGDT 285
QY 185 FFLYSLKTVDKSRWQQGNVFCSCVMHEALHNHYTKQSLSPGKDWLKAFYDKVAEKLKE 244
Db 286 YFLYSLKTVDTDSWLGQEIFTCVSVVHEALHNHHTQKNLSRSEPLE-----LNETCAE 337
QY 245 AFD-----WL 249
Db 338 AQDGLDGLWTITITIFISLFLSVCSASVTLFVKW 375

RESULT 9
GC3M_MOUSE STANDARD; PRT; 333 AA.
AC P20761;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2B chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;

```

```
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
DR PIR: PS0018; PS0018.
DR HSP: P01842; 7FAB.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_c1.
DR InterPro: IPR003600; Ig_like.
DR Pfam: PF00047; Ig; 3.
DR SMART: SM00410; Ig_like; 1.
DR SMART: SM00407; Ig_c1; 2.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 80
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 147 207
FT DISULFID 253 311
SQ SEQUENCE 333 AA; 36497 MW; 55F8B64D48D460A6 CRC64;

Query Match 58.7%; Score 840; DB 1; Length 333;
Best Local Similarity 62.1%; Pred. No. 1.4e-20;
Matches 151; Conservative 34; Mismatches 43; Indels 15; Gaps 2;

QY 1 MKT-----HTCP-----PCPAPELLGGPSVFLPPPKDITLMSRPEVTCVVV 45
DB 91 VDKVRRNGGIGKCPCTCHKCPVPLGGPSVFIIPPKDKTLLISQNAKVTCCV 150
QY 46 DYSHDEPVKNVYDGVVHNAKTPREEQYNSTYRVSVLTVLHQLDNLGCKEYCKV 105
DB 151 DVSEEDPVQFVFNWVHNAKTPREEQYNSTYRVSVLTVLHQLDNLGCKEYCKV 210
QY 106 NKALPAPIEKTISKAKGQPREPVYTLTPSRDELTKNOVSLCLVKGFPVSDIAVWESN 165
DB 211 NKALPAPIEKTISKPKGLVKRQVYVMGPPTEQLTEQVSLTCLISGFLPNDIGVETSN 270
QY 166 GQPNYKTPPVLSGDSGFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSL 225
DB 271 GHIEKNYKTEPVMSDGSFFMYSLNVERSRWDSRAPVCSVWVHGLNHHVKSISRP 330
QY 226 PGK 228
DB 331 PGK 333

RESULT 10
GC1M_MOUSE
ID GC1M_MOUSE STANDARD; PRT; 393 AA.
AC P01869;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ig gamma-1 chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=80045036; PubMed=115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
gamma 1 chain gene.";
RL Cell 18:559-568(1979).
[2]
RP SEQUENCE OF 323-393 FROM N.A.
RX MEDLINE=82197626; PubMed=6804950;
RA Tyler B.M., Cowman A.F., Gerondakis S.D., Adams J.M., Bernard O.;
```

```
RT "mRNA for surface immunoglobulin gamma chains encodes a highly
conserved transmembrane sequence and a 28-residue intracellular
domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2008-2012(1982).
[3]
RN SEQUENCE OF 323-366 FROM N.A.
RX MEDLINE=82115295; PubMed=6799207;
RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,
RA Eisenberg D., Wall R.;
RT "Gene segments encoding transmembrane carboxyl termini of
immunoglobulin gamma chains.";
RL Cell 26:19-27(1981).
[4]
RP SEQUENCE OF 1-44 FROM N.A.
RX MEDLINE=8222190; PubMed=6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
immunoglobulin gamma chains.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
CC -!- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA
SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
SEGMENT OF MU CHAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00793; CAA24172.1; -
DR EMBL; V00793; CAA24173.1; -
DR PIR; B02159; GLMSM.
DR HSP; P01842; 7FAB.
DR MGD; MGI:96446; Igh-4.
DR InterPro: IPR003006; IG_MHC.
DR InterPro: IPR003597; Ig_c1.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; Ig_c1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing; Transmembrane.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 110 HINGE.
FT DOMAIN 111 217 CH2.
FT DOMAIN 218 324 CH3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 138 198
FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .).
FT DISULFID 244 302 POTENTIAL.
FT TRANSMEM 340 357 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 358 393 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 393 AA; 43386 MW; 4CC88343B7AICE27 CRC64;

Query Match 57.8%; Score 827; DB 1; Length 393;
Best Local Similarity 51.7%; Pred. No. 6.4e-20;
Matches 149; Conservative 49; Mismatches 46; Indels 44; Gaps 6;

QY 7 CPP--CPAPELLGGPSVFLPPPKDITLMSRPEVTCVVVDSHEDPEVFNWVDSVGE 64
DB 104 CKPCICTVPEV---SSVFIFPPKPKDVLITLTLPKVTCTVVDISKDDPEVQSFVDDVE 160
QY 65 VHNARKTPREEQYNSTYRVSVLTVLHQLDNLGCKEYCKVSKVSNKALPAPIEKTISKAKGP 124
```


[illegible]

 Biochem. J. 126:837-850(1972).

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

```
DR GlycoSuiteDB; P01868;
DR MGD; MGI:96446; Igh-4.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_cl.
DR Pfam; PF00047; Iq; 3.
DR SMART; SM00407; Igc1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
FT Alternative splicing.
FT NON_TER 1
FT DOMAIN 1 97
FT DOMAIN 98 110
FT DOMAIN 111 217
FT DOMAIN 218 324
FT DISULFID 27 82
FT DISULFID 102 102
FT DISULFID 104 104
FT DISULFID 107 107
FT DISULFID 109 109
FT DISULFID 138 198
FT CARBOHYD 174 174
FT
FT
FT DISULFID 244 302
FT MOD_RES 324 324
FT CONFLICT 276 276
FT CONFLICT 278 278
SQ SEQUENCE 324 AA; 35704 MW; A338812F3D1F2C93 CRC64;

Query Match 57.6%; Score 823; DB 1; Length 324;
Best Local Similarity 51.3%; Pred. No. 4.8e-20;
Matches 142; Conservative 44; Mismatches 35; Indels 56; Gaps 3;

QY 5 HT-----CPP--CP 11
DB 51 HTEPAVLQSLDYLTLSSVTPSPRSEVTVCNVAHPASSTVKDKIVPRDCGCKPCIT 110
QY 12 APELLGGPSVFLFPPPKDLMISRTPEVTVVVDVSHEDPEVKFNWYVDGVEVHNAKTK 71
DB 111 VPEV---SSVFIPPKPKDVLITLTPKVTCTVVVDISKDPEVQFSWFVDVVEHTAQTO 167
QY 72 PREEQYNSTYRVSVLTVLHODWLNGLKEYCKVSKNKPALPAIEKTIKAKGP 131
DB 168 PREEQNFSTFRSVSELPIMHODWLNGLKEYCKVSKNKPALPAIEKTIKAKGP 227
QY 132 LPSPRDELTKNOVSLTCLVKGFPSPDIAYEVESNGQPNENYKTPPVLDSDGSEFLYSKL 191
DB 228 IPPPKQMAKDKVSLTCMTIDFEPDIIVENQWNGQPAENYKNTQPIIMNTNGSYFYVSKL 287
QY 192 TVDKSRWQGNVFCSCVMHEALHNHYTKLSLSLSPGK 228
DB 288 NVQKSNWEAGNFTCSVLHDEGLNHNHTEKLSLSPGK 324

RESULT 13
GCAM_MOUSE
ID GCAM_MOUSE STANDARD; PRT; 399 AA.
AC P01865;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2A chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82222190; PubMed=6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
RL immunoglobulin gamma chains";
RT Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
CC -!- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA
```

```
CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
CC BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
CC IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
CC SEGMENT OF MU CHAINS.
CC -!- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 1-329 IS ASSUMED TO BE
CC IDENTICAL WITH THE CORRESPONDING REGION OF THE SECRETED FORM OF
CC THE A ALLELE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J00471; AAB59661.1; ALT_INIT.
CC PIR; A02154; G2MSAM.
CC HSP; P01842; 7FAB.
CC MGD; MGI:96443; Igh-1.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR003597; Ig_cl.
CC Pfam; PF00047; Iq; 2.
CC SMART; SM00410; Igc1; 1.
CC SMART; SM00407; Igc1; 2.
CC PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 144 204
FT DISULFID 250 308
FT TRANSMEM 346 363
FT DOMAIN 364 399
FT CARBOHYD 180 180
FT SEQUENCE 399 AA; 44020 MW; 4C38138BFAED3FF0 CRC64;

Query Match 57.3%; Score 819; DB 1; Length 399;
Best Local Similarity 54.3%; Pred. No. 1.2e-19;
Matches 152; Conservative 33; Mismatches 52; Indels 43; Gaps 4;

QY 7 CPP--CPAPELLGGPSVFLFPPPKDLMISRTPEVTVVVDVSHEDPEVKFNWYVDGVE 64
DB 107 CPPCKCAPNLLGGPSVFIFPPKIKDVLMLSLSPITVTCVVVDVSDPDQVQISWFVNVE 166
QY 65 VNAKTKPREEQYNSTYRVSVLTVLHODWLNGLKEYCKVSKNKPALPAIEKTIKAKGP 124
DB 167 VTAQTQTHREDYNSTLRVVSALPIQHQDWMSGKFEKCKVNNKDLPAPIERTISKPGSV 226
QY 125 REPQVYTLPPSRDELTKNQVSLTCLVKGFPSPDIAYEVESNGQPNENYKTPPVLDSDGS 184
DB 227 RAPQVYVLPPEEEMTKKQVLTCTVDFMPEDIVVETNNGKTELNVKNTPEVLDSDGS 286
QY 185 FFLYSKLVKDVKSROOGNWFSCVMHEALHNHYTKLSLSLSPGK----- 229
DB 287 YFMYSKLVKRVKKNVVERNSYSCSVVHGLNHHHTKFSRTPLGLDLDVCAEAQDGLD 346
QY 230 -W-----LKAFFDKVAEKLKE 244
DB 347 LMTTITIFISLFLSVSVCSASVTLFKVKWIFSSVVE-LKQ 385

RESULT 14
GCC_RAT
ID GCC_RAT STANDARD; PRT; 329 AA.
AC P20762;
DT 01-FEB-1991 (Rel. 17, Created)
```

```

DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2C chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON NCBI_TaxID=10116;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=88166903; PubMed=3127222;
RA Brueggemann M., Delmastro-Galfre P., Waldmann H., Calabi F.;
RT "Sequence of a rat immunoglobulin gamma 2c heavy chain constant
RL Eur. J. Immunol. 18:317-319(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X07189; CAA30169.1;
DR PIR; S00847; S00847.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00410; IG-like; 1.
DR SMART; SM00407; IGcl; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 222 CH2.
FT DOMAIN 223 329 CH3.
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 113 113 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 143 203 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 249 307
SQ SEQUENCE 329 AA; 36571 MW; 5FCD7B7933850773 CRC64;

Query Match 57.2%; Score 818; DB 1; Length 329;
Best Local Similarity 57.9%; Pred. No. 7.3e-20;
Matches 146; Conservative 42; Mismatches 38; Indels 26; Gaps 2;

QY 3 KTHTC-----PP---CPAPELLGGPSVFLFPPKPKDTLMISR 36
Db 78 QTVCSVAHPATKSNLIKRIEPRRRPRPTDICSDDNLGRPSVFIFPPKPKDILMITL 137
QY 37 TPVETCVVDVSHEDPEVAFNVDGVVHNAKTKPREQYNSTYRVSVLTFLVHODWLN 96
Db 138 TPVETCVVDVSEEDVQFSEVDNVVFTVTAQTQHEQLNGTFRVSTLHLOHODWMS 197
QY 97 GKYEKCKVSNKALPAPIETKISKAKQREPPQVYTLPPSRDELTKNQVSLTCLVKGFYPS 156
Db 198 GKFEKCKVNNKOLPSPIETKISKPRKARTPVYTIPTPPREQMKNKSVLTCTVTSFYPA 257
QY 157 DIAVENESGCPENNYKTPPPVLDGSGFLYSLKLVVDKSRWQGNVFCSSVMEALHNH 216
Db 258 SISVERNGELSQDKNTLPVLDSDSYFLYSLKLVDSWNRGDIYTCSSVMEALHNH 317
QY 217 YTKSLSLSPGK 228
Db 318 HTOKNLSRSPGK 329
RESULT 15

```

```

GCAA_MOUSE STANDARD; PRT; 330 AA.
ID AC P01863;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ig gamma-2A chain C region, A allele.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=81076554; PubMed=6777755;
RA Sikorav J.-L., Auffray C., Rougeon F.;
RT "Structure of the constant and 3' untranslated regions of the murine
RL Balb/c gamma 2a heavy chain messenger RNA.";
RN [2]
RN Nucleic Acids Res. 8:3143-3155(1980).
RX SEQUENCE FROM N.A.
RX MEDLINE=81198976; PubMed=6262729;
RA Yamawaki-Kataoka Y., Miyata T., Honjo T.;
RT "The complete nucleotide sequence of mouse immunoglobulin gamma 2a gene
and evolution of heavy chain genes: further evidence for intervening
sequence-mediated domain transfer.";
RN [3]
RN Nucleic Acids Res. 9:1365-1381(1981).
RX SEQUENCE FROM N.A.
RX MEDLINE=81223894; PubMed=6787604;
RA Ollo R., Auffray C., Mochamps C., Rougeon F.;
RT "Comparison of mouse immunoglobulin gamma 2a and gamma 2b chain genes
suggests that exons can be exchanged between genes in a multigenic
family.";
RN [4]
RN Proc. Natl. Acad. Sci. U.S.A. 78:2442-2446(1981).
RX MYELOMA PROTEIN MOPC 173.
RX MEDLINE=74175517; PubMed=4831970;
RA Bourgois A., Fougereau M., Rocca-Serra J.;
RT "Determination of the primary structure of a mouse IgG2a
immunoglobulin:amino-acid sequence of the Fc fragment. Implications
for the evolution of immunoglobulin structure and function.";
RN [5]
RN Eur. J. Biochem. 43:423-435(1974).
RX DISULFIDE BONDS.
RX MEDLINE=73056887; PubMed=4565406;
RA de Preval C., Fougereau M.;
RT "Determination of the primary structure of a mouse gamma G2a
immunoglobulin. Identification of the disulfide bridges.";
RN Eur. J. Biochem. 30:452-462(1972).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00798; CAA24178.1;
DR PIR; A02152; G2MSA.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00410; IG-like; 1.
DR SMART; SM00407; IGcl; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107

```

```
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 144 204
FT DISULFID 250 308
FT MOD_RES 330 330 REMOVED POST-TRANSLATIONALLY.
SQ SEQUENCE 330 AA; 36389 MW; B84361C5445A6864 CRC64;

Query Match 56.9%; Score 813; DB 1; Length 330;
Best Local Similarity 52.9%; Pred. No. 1,1e-19;
Matches 148; Conservative 30; Mismatches 46; Indels 56; Gaps 2;

QY 5 HT-----CPP- 9
Db 51 HTFPVLOS DLYTLSSSVTVTSSTWPSQSI TCNVAHPASSTKV DKKIEPRGPTIKPCPC 110
QY 10 -CPAP ELLGGPSVFLFPKPKDTLMSRTPEVTCVVVDVSHEDPEVKENWYVDGVEVHNA 68
Db 111 KCPAPNLLGGPSVFIFFPKIKDVLMSLSPIVTCVVVDVSEDDPDVQISWFEVNNVEVHTA 170
QY 69 KTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIISKAKGQPREPQ 128
Db 171 QTQTHREDYNSTLRVVSALPIQHODWMSGKEPKCKVNNKDLPAPIERTISKPKGSVRAPQ 230
QY 129 VYTLPPSRDELTKNQVSLTCLVKGYGSPDIAVEWESNGQPENNYKTTPPVLDSDGSFELY 188
Db 231 VYVLPPEEEMTKQVTLTCWVTDPEMEDIYVETWNGKTELNYKNTPEVLDSDGSYFMY 290
QY 189 SKLTVDKSRWQGNVFCGSVMHEALHNYTKSLSLSPCK 228
Db 291 SKLRVEKKNNWERNYSYCSVWHEGLNHHHTTKFSRTPGK 330
```

Search completed: April 21, 2003, 10:43:13
Job time : 11.0941 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:21 ; Search time 38.8235 Seconds
(without alignments)
1401.120 Million cell updates/sec

Title: 2LINK7LINK7
Perfect score: 1430
Sequence: 1 MDKTHTCPPCPAPPELLGGPS.....AFDWLKAIFYDKVAEKLKEAF 264

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.0

Searched: 671580 seqs, 206047115 residues 671580
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_21:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phase:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1233	86.2	471	4	Q8TC77
2	1233	86.2	701	4	Q96FQ8
3	1143	79.9	473	4	Q8TC63
4	914	63.9	337	6	Q95M34
5	837	58.5	469	11	Q8R3V9
6	836	58.5	463	11	Q91C4
7	831	58.1	437	11	Q91A4
8	816	57.1	473	11	Q9D8L4
9	811	56.7	473	11	Q99L25
10	810	56.6	468	11	Q91L31
11	788	55.1	474	11	Q8R3H6
12	787	55.0	473	11	Q91Z05
13	513	35.9	34350	4	Q8WZ42
14	493	34.5	26926	4	Q10466
15	493	34.5	26926	4	Q8WZB3
16	440	30.8	17352	5	Q95YM2

ALIGNMENTS

RESULT 1

ID	Q8TC77	PRELIMINARY;	PRU;	471 AA.
AC	O8PC77;			
DT	01-JUN-2002 (TremBrel. 21, Created)			
DT	01-JUN-2002 (TremBrel. 21, Last sequence update)			
DE	01-JUN-2002 (TremBrel. 21, Last annotation update)			
DE	Hypothetical 51.8 kDa protein.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=SPLEEN;			
RA	Strausberg R.;			
RL	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; BC024289; AAH24289.1; -			
KW	Hypothetical protein.			
SQ	SEQUENCE 471 AA; 388F7F4CF588660E CRC64;			

Query Match	86.2%	Score 1233;	DB 4;	Length 471;
Best Local Similarity	100.0%;	Pred. NO. 6.5e-39;		
Matches 227;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	2	DKTHTCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD	61	
Db	245	DKTHTCPPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD	304	
QY	62	GVEVHNKTKPREQYNTYRVSVLTVLHQDLNGLNGKEYCKVSNKALPAPIETISKAK	121	
Db	305	GVEVHNKTKPREQYNTYRVSVLTVLHQDLNGLNGKEYCKVSNKALPAPIETISKAK	364	
QY	122	GQPREPQYVTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDS	181	
Db	365	GQPREPQYVTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDS	424	
QY	182	DGSEFLYSKLVNDKSRWQOGNWFSCVMHEALHNHYTOKSLSPGK	228	

```
Db 425 DGSFFLYSLKTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPGK 471
RESULT 2
Q96PQ8 PRELIMINARY; PRT; 701 AA.
AC Q96PQ8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Factor VII active site mutant Immunocjugate.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21477448; PubMed=11593034;
RA Hu Z., Garen A.;
RT "Targeting tissue factor on tumor vascular endothelial cells and tumor
RL cells for immunotherapy in mouse models of prostatic cancer.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12180-12185(2001).
DR EMBL: AF272774; AAK58686.1; -.
DR InterPro: IPR001152; ASX_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001254; Ser.protease_Try.
DR InterPro: IPR000294; VitK_dep_GLA.
DR Pfam: PF00008; EGF; 2.
DR Pfam: PF000594; gla; 1.
DR Pfam: PF00047; ig; 2.
DR Pfam: PF00089; trypsin; 1.
DR SMART: SM00181; EGF; 2.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
DR PROSITE: PS01187; EGF_Ca; UNKNOWN_1.
DR PROSITE: PS00011; GLU CARBOXYLATION; UNKNOWN_1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
DR PROSITE: PS02040; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 701 AA; 77826 MW; 94AC6CEB42CC992F CRC64;

Query Match 86.2%; Score 1233; DB 4; Length 701;
Best Local Similarity 100.0%; Pred. No. 2.4e-38;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DKTHTCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
Db 475 DKTHTCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 534
QY 62 GVEVHNAKTPREEQNSTYRVSIVLTLHQDLNGKEYCKKYSKALPAPIEKTISKAK 121
Db 535 GVEVHNAKTPREEQNSTYRVSIVLTLHQDLNGKEYCKKYSKALPAPIEKTISKAK 594
QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 181
Db 595 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 654
QY 182 DGSFFLYSLKTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPGK 228
Db 655 DGSFFLYSLKTVDKSRWQGNVFCSCVMHEALHNHYTKQSLSLSPGK 701

RESULT 3
Q8TC63
ID Q8TC63 PRELIMINARY; PRT; 473 AA.
AC Q8TC63;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)

Query Match 79.9%; Score 1143; DB 4; Length 473;
Best Local Similarity 84.7%; Pred. No. 2.1e-35;
Matches 211; Conservative 9; Mismatches 6; Indels 23; Gaps 1;

QY 3 KTHT-----CPPCPAPELGGPSVFLFPPKPKDTLMISRTPE 39
Db 225 KYTCNVNDHKPSNTKVDKRVESKYGPPCPCPAPELGGPSVFLFPPKPKDTLMISRTPE 284
QY 40 VTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTPREEQNSTYRVSIVLTLHQDLNGKE 99
Db 285 VTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTPREEQNSTYRVSIVLTLHQDLNGKE 344
QY 100 YCKVSKNKKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIA 159
Db 345 YCKVSKNKKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIA 404
QY 160 VEWESNGQPENNYKTPPPVLDSDGSFFLYSLKTVDKSRWQGNVFCSCVMHEALHNHYTQ 219
Db 405 VEWESNGQPENNYKTPPPVLDSDGSFFLYSLKTVDKSRWQGNVFCSCVMHEALHNHYTQ 464
QY 220 KSLSLSPGK 228
Db 465 KSLSLSLGK 473

RESULT 4
Q95M34 PRELIMINARY; PRT; 337 AA.
AC Q95M34;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Immunoglobulin gamma 1 heavy chain constant region
DE (Fragment).
DE IGHC1.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RA Wagner B.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98383416; PubMed=9717671;
RA Wagner B., Overesch G., Sheoran A., Holmes M., Richards C.,
RA Leibold W., Radbruch A.;
RT "Organization of the equine immunoglobulin heavy chain constant region
RL genes. III. Alignment of c-mu, c-gamma, c-epsilon and c-alpha genes.";
RL Immunobiology 199:105-119(1998).
DR EMBL: AJ300675; CAC44624.1; -.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig; 2.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_2.
FT NON_TER 1
SQ SEQUENCE 337 AA; 37438 MW; A60BF2B01DEFD1F6 CRC64;
```



```

RT antibody (Wab 7, its light and heavy chains) and construction of a
RL single chain antibody (scFv).";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF152372; AAD40243.1; -.
DR HSSP; P01842; 7FAB.
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR003600; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00406; Igv; 1.
DR SMART; SM00410; Ig-like; 2.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
FT NON_TER 1
FT TER 437
SQ SEQUENCE 437 AA; 48142 MW; 5C3A7BB3EE7D697C CRC64;

Query Match 58.1%; Score 831; DB 11; Length 437;
Best Local Similarity 51.6%; Pred. No. 2.2e-23;
Matches 143; Conservative 43; Mismatches 35; Indels 56; Gaps 3;

Oy 5 HT-----CPP--CP 11
||
Db 164 HTFPAVLQSLDYLTLSSVTPSPSTVTCNVAHPASSTKVDRKIVPRDCGCKPCICT 223
Oy 12 APELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDGVEVHNAKTK 71
||
Db 224 VPEV---SSVFPPKPKDVLITLTPKVTCTVVDVSHEDPEVQSFVDFVDTAQIQ 280
Oy 72 PREQYNSTRYRVSVLTVLHQDLNKGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVY 131
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 281 PREQFNSTRYRSVSELPIMHQDLNKGKEYKCKRVNSAAFPAPIEKTISKTKGRKAPQVY 340
Oy 132 LPSPRDELTKNOVSLTCLVKGYFSPVDIAVEWESNGOPENNYKTTTPVLDSDGSFLYSKL 191
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 341 IPPPEQMAKDKVSLTCMTIDFEPDITVEWQMGQPAENYKNTQPIMDTDGYSFYVSKL 400
Oy 192 TVDKSRWQOGNFGFSCVMHEALHNHYTQKSLSLSPGK 228
|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
Db 401 NVQKSNWEAGNTFTCSVLHLEGLHNHTKLSLSPGK 437

RESULT 8
Q9DBL4
ID Q9DBL4 PRELIMINARY; PRT; 473 AA.
AC Q9DBL4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 1810060009rik protein.
GN IGH-1 OR 1810060009RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

```

```

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK007918; BAB25349.1; -.
DR HSSP; P01842; 7FAB.
DR MGD; MGI:96443; Igh-1.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003600; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00409; Ig; 2.
DR SMART; SM00407; Igc1; 3.
DR SMART; SM00406; Igv; 1.
DR SMART; SM00410; Ig-like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 473 AA; 51699 MW; 9DED57A514475FBB CRC64;

Query Match 57.1%; Score 816; DB 11; Length 473;
Best Local Similarity 50.5%; Pred. No. 1.1e-22;
Matches 144; Conservative 38; Mismatches 42; Indels 61; Gaps 1;

Oy 5 HT-----
||
Db 189 HTFPALQSLGTYLTLSSVTVTSNTWPSQTITCNVAHPASSTKVDRKIVPRVITQNP 248
Oy 7 ---CPPCAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVDG 63
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 249 LKECPCAAPDLGGPSVFIPPKIKDVLMLSLSPWVTCVVDVSEDDPDVQISFVNVN 308
Oy 64 EVHNAKTRPEEQYNSTRYRVSVLTVLHQDLNKGKEYKCKVSNKALPAPIEKTISKAKG 123
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 309 EVHTAQQTTHREDYNTSLRVVSALPIQHDWMSGKEFKCKVNNRALPSPIEKTISKPRG 368
Oy 124 PREQVYTLPPSRDELTKNOVSLTCLVKGYFSPVDIAVEWESNGOPENNYKTTTPVLDSDG 183
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 369 VRAPQVYVLPFAEEETKRFSLTCMTITGFLPAEIAVDWNTSGRTPEQYKNTATVLDSDG 428
Oy 184 SFFLYSKLTVDKSRWQOGNFGFSCVMHEALHNHYTQKSLSLSPGK 228
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 429 SYFWYSLKRVQKSTWGRSLFACSVVHGLNHLNLTTKTISRSLGK 473

RESULT 9
Q99L25
ID Q99L25 PRELIMINARY; PRT; 473 AA.
AC Q99L25;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to RIKEN cDNA 1810060009 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEB-2001; to the EMBL/GenBank/DBJ databases.
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003888; AAH03888.1; -.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003600; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; Igc1; 3.

```



```

RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010327; AAH10327.1;
DR MGI; MGI:2144967; AU044919.
DR InterPro; IPR000345; Cytochrome_bbind.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_3.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Hypothetical protein.
SQ
SEQUENCE 473 AA; 51946 MW; CF625F008932AF12 CRC64;

Query Match          55.0%; Score 787; DB 11; Length 473;
Best Local Similarity 50.0%; Pred. No. 1.5e-21;
Matches 143; Conservative 34; Mismatches 47; Indels 62; Gaps 2;

QY 5 HT-----CP-----CP-----CP-----CP-----CP 8
II
Db 188 HTFPALLOGLYTMSSSVTPSPSTWPSQVTGSAHPASSTVTDKKLEPSGPISTINPCP 247
II
QY 9 P-----CAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDG 62
II
Db 248 PKCEKCHKAPNLEGGPSVFIFPNIKDVLMLSLTPKVTCTVVVDVSEDDPDVQISWVNN 307
II
QY 63 VEVHNAKTPREEQVNSTYRVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIKSKAG 122
II
Db 308 VEVHTAQTQTHREDYNSTIRVVSALPIQHQDAMSGCKEPCVKNNDLPSPIERTISKIG 367
II
QY 123 QREPOVYTLPSRDELTKQVSLCLVKGFPDSDIAVWESNGQFPENNYKTPPVLDSD 182
II
Db 368 LVRAPOVYTLPPAEQLSRKDSLTCLVVGFPNGDISVWTSNGHTENYKDTAPVLDSD 427
II
QY 183 GSFYLSKLTVDKSRWQGNVFCSCVMHEALNHVYTKSLSLSPGK 228
II
Db 428 GSYFIYKLDITKWEKTDSDSCNVRHGLKNTYLLKTIISRSPGK 473
II

RESULT 13
Q8WZ42
ID Q8WZ42 PRELIMINARY; PRT: 34350 AA.
AC Q8WZ42;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Titin.
GN TTN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RX MEDLINE=20309627; PubMed=10850961;
RA Freiburg A., Trombitas K., Hell W., Cazorla O., Fougereousse F.,
RA Cencner T., Kolmerer B., Witt C., Beckmann J.S., Gregorio C.C.,
RA Granzier H., Labeit S.;
RT "Series of exon-skipping events in the elastic spring region of titin
RT as the structural basis for myofibrillar elastic diversity.";
RL Circ. Res. 86:1114-1121(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21573839; PubMed=11717165;
RA Bang M.L., Centner T., Fornoff F., Geach A.J., Gotthardt M.,
RA McNabb M., Witt C.C., Labeit D., Gregorio C.C., Granzier H.,
RA Labeit S.;
RT "The complete gene sequence of titin, expression of an unusual ~700
RT kDa titin isoform and its interaction with obscurin identify a novel
RT Z-line to I-band linking system.";
RL Circ. Res. 89:1065-1072(2001).
DR EMBL; AJ277892; CAD12456.1;
DR InterPro; IPR000282; Cytok_receptor_2.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR000577; FG_Y_kin.

```

```

DR InterPro; IPR003961; FN_III.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR InterPro; IPR002016; Peroxidase.
DR InterPro; IPR004168; PPAK_motif.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00041; fn3; 132.
DR Pfam; PF00047; Ig; 146.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF02818; PPAK; 53.
DR ProDom; PM000001; Euk_pkinase; 1.
DR SMART; SM00080; FN3; 133.
DR SMART; SM00409; IG; 167.
DR SMART; SM00408; IGC2; 148.
DR SMART; SM00406; IGV; 23.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; Tyrcg; 1.
DR PROSITE; PS00933; FG_Y_KINASES_1; UNKNOWN_1.
DR PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
SQ SEQUENCE 34350 AA; 3816262 MW; 5B1120058A7CE58A CRC64;

Query Match          35.9%; Score 513; DB 4; Length 34350;
Best Local Similarity 0.5%; Pred. No. 7.4e-05;
Matches 172; Conservative 36; Mismatches 54; Indels 32039; Gaps 47;

QY 3 KTHTCPP-----9
II
Db 98 KAETAPPNFVRLQSMVTVRQGSQVRLQVRVTGIPTPVVKYRDGAIEQSSLDFOISQEGD 157
II
QY 10 -----9
Db 158 LYSLLIAEAYPEDSGTYSVNATNSVGRATSTAELLVQGEVEVPAAKTKTIYSTAQSISR 217
QY 10 -----9
Db 218 QTRIEKKIEAHFDARSATVEMVIDGAAGQQLPKHTPHRIPPKPKRSRPTPPSIAKAQL 277
QY 10 -----9
Db 278 ARQSPSPIRHSPSPVRHVRAPTPSPVRSVSPAARISTSPRSPLLMRKTKQASTVAT 337
QY 10 -----9
Db 338 GPEVPPPKQEGYVASSSEAEEMRETTLTSTQIRTEWEGRYGVQEQVTISGAAGAAAS 397
QY 10 -----9
Db 398 VSASASAAEAATGATKEVKQDADKSAATVVAADVMARVREPVISAEQTAQTITTTA 457
QY 10 -----9
Db 458 VHIQPAQEQVRKEAKTAVTKVVVAADKAKEQLKSRTKEVITTKQEQMHVTHEQIRKET 517
QY 10 -----9
Db 518 EKTFFVKVVISAAKAKEQETRISEITKKQKVQTEAIRQETITEAASVWVVAATAKTKL 577
QY 10 -----9
Db 578 ETVPGAQEBETTQDDQMHLSEYKIMKTRKTVVVKVIVATPKVREQLVSRGREGITTKR 637
QY 10 -----9
Db 638 EQVQITQEKMRKEAKETALTSTIAVATAKAKEQETILTRTRETMTATREQEIQVTHGKVDVGK 697

```

QY 10 ----- 9
Db 698 KAEAVTVAADQAVREPREPHLEESVAQOTTLEYGYKERISAAKVAEPPQRPASEP 757
QY 10 ----- 9
Db 758 HVPKAVKPRVIOAPSETHIKTTQKGMHISSOIKKTTDLTTERLVHVDKRPRTASPHFT 817
QY 10 ----- 9
Db 818 VKSISVPKTEHGEYASAGSAIATLOKELSATSSAQKITKSVKAPVKPSETRVRAEPTP 877
QY 10 ----- CPAP-- 13
Db 878 LPOFPADTPTYKSEAGVEVKEVGSITGTTVREERFEVLHGREAKVTETARVPAPVE 937
QY 14 ----- ELGGPSYFL-----FP----- 25
Db 938 IPVTPPTLVGLKNVTIEGESVTLECHISGYPSPTVTWYREDYQIESSIDFOITFQSGI 997
QY 26 ----- 25
Db 998 ARLMIREAFADSGRTCSAVNEAGIVSTSCYLAVQVSEFEKETTAVTEKFTTEKRFV 1057
QY 26 ----- PKP----- 28
Db 1058 ESRDVVMTDTSLEQAGPGEAPAPVFTKPVVOKLVGGSVVFGQVGNPKPHVYWK 1117
QY 29 ----- 28
Db 1118 SGVPLTGYRYKVSYNKQTECKLVISMFTFADDAGEYTIIVRNKHGETSASALLEEADY 1177
QY 29 ----- KDTLM-- 33
Db 1178 ELLMKSOEMLYQVTAFOEPKVGETAPGFVYSEYEKEQALIRKKMAKDIVVVR 1237
QY 34 ----- 33
Db 1238 TYVEDQEFHISPEERLIKEIYRIIKTTLEELLEEDGEKMAVDISEASEVSGFDLRI 1297
QY 34 ----- 33
Db 1298 KNYRILEGMGVTHCKMSGYPLPKIAWKDGKRIKHGERYQMDFLODGRASLRIPVVLPE 1357
QY 34 ----- ISRTP-- 38
Db 1358 DEGIYTAASNIGNAICSGKLYVEPAAPLGAFTYIPTLEPVSRIRSLSPRSVSRPIRM 1417
QY 39 ----- 38
Db 1418 SPARMSPARMSPARMSPARMSPGRRLLEETDESOLERLYKPVFVLKPVSKFCLEGQTARFD 1477
QY 39 ----- 38
Db 1478 LKVYGRMPETFWHDGQQIVNDYTHKVVIKEDGTQSLIIVPATPSDGEWTVVAQNRA 1537
QY 39 ----- 38
Db 1538 RSSISVILTVEAVHQVPMFVEKLNKVNIEKGSQLEMKVRCATGNPNPDIVWLKNSDIIV 1597
QY 39 ----- 38
Db 1598 PHYPKIRIBGTGEAALKIDSTVSQDSAWYTATAINKAGRDYTRCKVNVVEFEAPEPE 1657
QY 39 ----- 38
Db 1658 RKLIIIPRGYRAKEIAAEPLHLRYGOEWEEDLYDKEKQKPPFKKLTSLRKRF 1717
QY 39 ----- EUTC 42
Db 1718 GPAHFECRLTPIGDTMVVWHLHDGKPLEANRLRMINIEFGYCSLDYGVAYSRDSGIITC 1777
QY 43 ----- 42

Db 1778 RATNKYGTDHTSATLIVKDEKSLVEESQLEPRKGLQRIEBELERMAHEGALTGVTTDOKE 1837
QY 43 ----- 42
Db 1838 KQKPDIVLYPEPVRLVEGETARFCRVGTGYQPKVNMWYLNQGLIRKSRFRVYDGIHYL 1897
QY 43 ----- 42
Db 1898 DIVDCKSYDTGEVYKTAENPEGVIEHKVKLEIQOQREDFRSVLRAPRPEPFHVHPEGKL 1957
QY 43 ----- 42
Db 1958 QFEVQKVDRPVDTTETKEVVVKLRAERITHEKVPSESEELRSKFRKRTTEGYEAITAVE 2017
QY 43 ----- 42
Db 2018 LKSRKXDESYEELLRTKDELLHWTKELTTEEXKALAEKGKITPTPKPKIELSPSMEA 2077
QY 43 ----- V 43
Db 2078 PKIFERIOSQTVGGGSDAHEFRVRVVGKPDPECEWYKNGVKIERSDRIYVWPEDNVCELV 2137
QY 44 VVDVSHED----- 51
Db 2138 IRDVTAEADSASIMVKAINIAGETSSHAFLVLQAKQLITFTQELQDVVAKEDTMAFECE 2197
QY 52 ---PEVKFNWYDGVVH-----NAKT- 70
Db 2198 TSEPFVKVWKYKDGMEVHEGDKYRMHSDRKVHFLSILTDTSADYDSCVLVEDENVKTT 2257
QY 71 ----- 70
Db 2258 AKLIVEGANVEFVKELQDIEVPESYSGELECIYSPENIEGKWHNDVELKSNKYTTISR 2317
QY 71 ----- KPR----- 73
Db 2318 RGRQNLTVKDVTKEDQGEYSFVIDGKTTCKLKMKPRPIALQGLSDQKVCBGDIVQLEV 2377
QY 74 ----- 73
Db 2378 KVSLESVEGVMKDGQEVQPSDRVHIVIDKOSHMLLIEDMTKEDAGNYSFTIPALGLSTS 2437
QY 74 ----- 73
Db 2438 GRVSVSYVDVITPLKDVNVIEGTAVLECKVSVDPVTSVKWYLNDEQIKPDDRQVAIVKG 2497
QY 74 ----- 73
Db 2498 TKQRLVINRTHASDEGPYKLIIVGRVETNCNLSVEKIKIIRGLRDLTCTETQNVVFEVELS 2557
QY 74 ----- 73
Db 2558 HSGIDVLWNFKDKRKPSKSKYKIEAHGKIYKLTVLNMKDDKGYTFYAGENITSGKLTV 2617
QY 74 ----- 73
Db 2618 AGGAISKPLTDQTVAESQEAFFECEVANPDSKGEWLBDGKHLPLTNINRSESOGHKRRLI 2677
QY 74 ----- 73
Db 2678 IAATKLDDIGEYTYKVATSKTSKAKLKEAVKIKKTLKNLTVTETQDAVFTVELTHPNVKG 2737
QY 74 ----- 73
Db 2738 VOWIKNGVLESNEKYAISVKGTIYSLRIKNCIAIVDESIVGFRGLRGLSARLHVETVKI 2797
QY 74 ----- 73
Db 2798 IKPKDVTALENATVAFESVSHDTPVVKWFHKNVEIKPSDKHRLVSEKRVHKLMLQNIS 2857
QY 74 ----- 73

Db 2858 PSDAGEYTA VVGQLECKAKLFVETLHITKMTKNIEVPETKTASFCEVSHFNVPMSMLKN 2917
Qy 74 ----- 73
Db 2918 GVEIEMSEKFI VVGQKLHQLIIMNTSTEDSAEYTFVCGNDQVSATLVTPIIMITSMKLD 2977
Qy 74 ----- 73
Db 2978 INAEKDIRITFEVTYNEGISYKWLKNGVEIKSTDKCOMRTKLLTHSLNIRNVHFGDAAD 3037
Qy 74 ----- 73
Db 3038 YTFVAGKATSTATLYEARHIEFRKHIDIKVLEKKRAMFECEVSEPDITVQMMKDDQEL 3097
Qy 74 ----- 73
Db 3098 QITDRIKIQEKYVHRLIIPSTRMSDAGKYTVVAGNVSTAKL FVEGRDVRIRSIKKEVO 3157
Qy 74 ----- 73
Db 3158 VIEKQRAVVEFEVNEDDVAHWYKDGIEINFQVQERHKYVVERRIHRMFISETRQSDAGE 3217
Qy 74 ----- -BEQ 76
Db 3218 YTFVAGNRSSVTLXVNAPEPQVLQELQPVTVQSGKPARFCAVISGRPQPKISWYKEEQ 3277
Qy 77 YNST----- 80
Db 3278 LLSTGFKCKFLHDGQEYTLILLIEAFPDAANYTCEAKNDYGVATTSASLSVEPVSFDP 3337
Qy 81 ----- 80
Db 3338 QEMPVYPAIITPLQDITVTSEGPAPFQCRVSGTDLKVSWYSKDKKIKPSRFRMTQPED 3397
Qy 81 ----- 80
Db 3398 TYQLEIAEYPEDECTYTFVNASNAVQVSVSTANLSLEAPESILHERIQEIEMEMKEFSS 3457
Qy 81 ----- 80
Db 3458 SFLSAEEBGLHSAELQLSKINETLLESPPYSTKFDSEKEGTCPIFIKEVSNADISMG 3517
Qy 81 ----- 80
Db 3518 DVATLSVTVIGIPKPIQWFFNGVLLTPSADYKFVFDGDDHSLIILFTKLEDEGEYTCMA 3577
Qy 81 ----- 80
Db 3578 SNDYKGTICSAYLKINSKGEGHKDTETESAVAKSLEKLGGPCPPHFLKELKPIRCAOGLP 3637
Qy 81 ----- 80
Db 3638 AIFEYTVVGEPAPTVTFWKENQOLCTSVYTIHNPNGSGTFIVNDPQREDSGLYICKAE 3697
Qy 81 ----- 80
Db 3698 NMLGESTCAELLVLEDTMDTDPCKAKSTPEAPDPPTPLKGPAAVEALDSEQIATF 3757
Qy 81 ----- 80
Db 3758 VKDTILKAALITEENQQLSYEHIAKANELSSQLPLGAQELQSLQEDKLTPESTREFLCI 3817
Qy 81 ----- 80
Db 3818 NGSIHFQPLKEPSPNLQIIVQSQKTFSEKIGILMPEEPETQAVLSDTEKIFPSAMSIEQI 3877
Qy 81 ----- 80
Db 3878 NSLTVPLKTLAEPENYPOSSIEPPMHYSYLTSAVEEVLSPKEKTVSDTNREQRVTLQK 3937
Qy 81 ----- 80
Db 3938 QEAQALISLSLAEGHVESLQSPDVMISQVNYEPLVPSEHSCFEGGKILIESANPLENA 3997

Qy 81 ----- 80
Db 3998 QODSAVRIEKGSLRFPPLALEEKQVLLKEEHSNDNVMPDPQIIIESKREPVAIKKVQEVQG 4057
Qy 81 ----- 80
Db 4058 RDLLSKESLLSGIPEQRLNLKIQICRALQAASVASEQPLFSEWLNRIEKVEVEAVNITQ 4117
Qy 81 ----- 80
Db 4118 EPRHIMCYLVTSAKSVEEVTIIIEDVDPOMANLKMELRDALCAIIEEIDILTAEGPR 4177
Qy 81 ----- 80
Db 4178 IQQAKTSLQEMDSFSGSKQKVEPITEPEVESKYLSPEEVSYNVQSRVKYLDATPVTK 4237
Qy 81 ----- 80
Db 4238 GVASAVSDEKQESLKPSEKESSSESGTEEVATVKIQEAGGFIKEDGPMIHTPLVD 4297
Qy 81 ----- 80
Db 4298 TVSEEGDIVHLTTSITNAKEVNWYFENKLVPSDEKFKCLQDONTYTLVIDKVNTEHDQGE 4357
Qy 81 ----- 80
Db 4358 YVCEALNDSGKTATSAKLTVVKRAAPVIKRIEPLVALGHAKFTCEIQSNAPNVRFOWF 4417
Qy 81 ----- 80
Db 4418 KAGREIVESDKCSIRSSKYISSLEILRTQVVVDCGEYTCKASNEYSVSCTATLTVTAEYP 4477
Qy 81 ----- 80
Db 4478 PTFLSRPKSLTTFVGKAAKFTCTVTGTPTVIETIWOKGAALSPSPNWKISDAENKHILEL 4537
Qy 81 ----- 80
Db 4538 SNLTIQDRGVYSCASNKFGADICQAEI IIDKPHFIKEPQVQSAINKKVHLECOVDDE 4597
Qy 81 ----- 80
Db 4598 RKVTWTWSKDGKQLPPGKDYKICFEDKTIATLEIPLAKLDSGTVYCTASNEAGSSCSAT 4657
Qy 81 ----- 80
Db 4658 VTVREPPSFVKKVDPSYLMPLPGESARLHCKLKGSPVIQVTFWKNKKELSSENTVRMYEYN 4717
Qy 81 ----- 80
Db 4718 SEAILDITDKVEDSGSYSCAVNDVGDSCSTEIVIKEPPSFIKTLEPADIVRGTNALL 4777
Qy 81 ----- 80
Db 4778 QCEYSGTGPEPISWFKDKQIRSSKKYRLFQSKSLVCLFISFNSADVGEYCVVANEVG 4837
Qy 81 ----- 80
Db 4838 KCGCMATHLLKEPPTFVKVDDLLIAGGQVTLQAAVRGSEPISVTMWKGQEVIREDGKI 4897
Qy 81 ----- 80
Db 4898 KMSFSGVAVILLIPDVQISFGGKYTKLAENEAGSQTSGVGLIVKEPAKIIERAELIOVTA 4957
Qy 81 ----- 80
Db 4958 GDPATLEYTVAGTPELKPWKYKGRPLVASKKYRISFPKNVNAQLKFFYSAELHDSQCTFE 5017
Qy 81 ----- 80
Db 5018 ISNEVGSSCETTFTVLDRDIAPFTTKPLRNVDSVWNGTCRLDCKIAGSLPMRVSFWDG 5077

QY	81	-----80
Db	5078	KEAASDRYRIAFVEGTASLEIIRVDMNDAGNETCRATNSVGSKDSSGALIVQEPSPFT 5137
QY	81	-----80
Db	5138	KPGSKDVLPGSACLKSTFGOSTPLTIRWPKGNKELVSGSCYITKEALESSELYLVKT 5197
QY	81	-----80
Db	5198	SDSGTYTCKYSNVAGGYECSANLFFVKEPATFVEKLEPSQLLKKGDATQLACKVTGTPPIK 5257
QY	81	-----80
Db	5258	ITWFANDREIKESKHRMSFVESTAVLRLTDVGIEDSGEYMCBAQNEAGSDHCSIVIVK 5317
QY	81	-----80
Db	5318	ESPYFTKEFKPIEVLKEYDVMLLAEVAGTPPEITWFKDNTILRSGRKYKTFIODHLVSL 5377
QY	81	-----80
Db	5378	QILKFVAADAGEYQCRVTNEVGSSICARSRTLREPPSFIIKIBSTSSLRGCTAAFAQTLK 5437
QY	81	-----80
Db	5438	GSLPITVTWLKDSDEITEDDNIIRMTFENNVAASLYLSGIEVKHDGKYVCAKNDAGIQRCS 5497
QY	81	-----80
Db	5498	ALLSVKEPATITEEAVSIDVTQGDPAFLQVFKFSGTKEITAKWFKDGOELTLGSKYKISVT 5557
QY	81	-----80
Db	5558	DTVSIILKIIISTEKKDSGEYTFEVDNDVGRSCKARINVLIIIPPSFTKKLKMDSIKGS 5617
QY	81	-----80
Db	5618	FIDLECIVAGSHPISIQWKDDQEIASEKYKFSFHONTAFLEISOLEGTDGTYTCSAT 5677
QY	81	-----80
Db	5678	NKAGHNOCSGHLTVKEPPYFEVERQSDVNPNTRVOLKALVGGTAPMTIKWFKDNKELHS 5737
QY	81	-----80
Db	5738	GAARSVKDDTSTSLFPAKATDSGTYICQLSNDVGTATSKATLFFVKEPPQFIKKPSPV 5797
QY	81	-----80
Db	5798	LVLNRGOSTTFEQITGTPKIRVSWYLDGNEITAIQKHGISFIDGLATFOISGARVENS 5857
QY	81	-----80
Db	5858	TYVCEARNAGTASCISIELKVKEPPTFIRELKPVEVVKYSDVELECEVTGTPPEFTWLK 5917
QY	81	-----80
Db	5918	NNREIRSKKYTLTDRVSVFNLHITKCDPSDTGEYQCIIVSNEGSCSCTRVALKEPPSF 5977
QY	81	-----80
Db	5978	IKKIENTTVLKSSATFQSTVAGSPPISTWLKDDQILDEDDNVIISFVDSVATLQIRSV 6037
QY	81	-----80
Db	6038	DNGHSGRYTCAKNESGVERCYAFLLVQEPQAIVERAKSVDTVKDPMTECVVAGTPEL 6097
QY	81	-----80
Db	6098	KVKWLKDGKQIVPSRYSFMSFENNVAFRIOQVWVKQDSQYTFKVENDFGSSCDAYLRV 6157
QY	81	-----YRVV---84

Db	6158	LDQNIPPSFTKLTMKDKVLGSSIHMECKYSGSLPISAOWFKDGKEISTSAKYRLVCHER 6217
QY	85	-----84
Db	6218	SVSLEVNMLEDTANTYTKVSNVAGDDACSGILTVKEPPSFLVKPGRQQAIPDSTVEFK 6277
QY	85	-----84
Db	6278	AILKGTPPFKIKWPKDDDELVSQPKCFIGLEGSTSFNLNLYSVDASKTCQYTCYHTNDVGS 6337
QY	85	-----84
Db	6338	DSCTTMLLVTEPPKFKVKLEASKIVKAGDSRLECKIAGSPEIRVWFRNEHELPAISKY 6397
QY	85	-----84
Db	6398	RMTFIDSVAVIQNMNLTSTEDSGDFICEAQNPAGSTSKVIVKEPPVFSFPPIVETLK 6457
QY	85	-----84
Db	6458	NAEVSLECELSGTPPPFVWVKDKRQLRSSKKYKIASKNPHTSIHLNVDTSIDIGEYHCK 6517
QY	85	-----84
Db	6518	AQNEVGSDTCVCTVKLKEPPRFVSKLSLTVVAGEPAELQASIEGAQPIFVQWLKEKEEV 6577
QY	85	-----84
Db	6578	IRESENIRITFVENVATLOFAKAPANAGRYICOIKNDGGMEENMATLMVLEPAVIVEKA 6637
QY	85	-----84
Db	6638	GPMTVTVGETCTLECKVAGTPELSVEWYKDGKLLTSSQKHKFSYKLISSLRILSVERQD 6697
QY	85	-----84
Db	6698	AGTVTFQVQNNVGKSCTAVVDSDRAVPPSFTRRLLKNTGGVLGASCILECKVAGSSPIS 6757
QY	85	-----SVLTV- 89
Db	6758	VAMFHEKTIKISGAKYQTTTFSDNVCTIQLNSLSDSDMGNTCTVAANVAGSDECRAVLTQ 6817
QY	90	-----89
Db	6818	EPFSFVKEPEPLEVLPCKNVTFTSVIRGTPPFKVNWFRGARELVKGRDRCNIYFEDTVAEL 6877
QY	90	-----89
Db	6878	ELFNIDISQGEYTCVYVNNAGQASCTTRLFVKPEAAFLKRLSDHSVEPGKSIILESTYT 6937
QY	90	-----89
Db	6938	GTLPISVTKKDGFNITTEKCNIVTTEKTCILEILNSTRKDRAGQYSCIEIENEAGRDVCG 6997
QY	90	-----89
Db	6998	ALYSTLEPPYFVTEPLEAARAVGDSVSLQCOVAGTPEITVSWYKGDTKLRTPPEYRYFT 7057
QY	90	-----89
Db	7058	NNVATLVFNKVINDSGEYTCKAENSIGTASCTVFRIOERQLPPSFARQLKDIEQTUGL 7117
QY	90	-----89
Db	7118	PVTLTCLRLNGSAPIQVCWYRDGVLLRRDENLQTSFVDNVATLKLQTLDSHSGOYSCSAS 7177
QY	90	-----89
Db	7178	NPLGTASSARLTAREPKSPFFDIKPVSIDVIAGESADFECHVTGAQPMRITWSKDNKE 7237
QY	90	-----89

Db 7238 IRPGNYTITCVGNTPLHRLKLVKGDGQYTCQATNDYKDMCSAQLSVKBPPEVKKL 7297
QY 90 -----LHODW----- 94
Db 7298 EASKVAKGESIQLECKISGSPKIVSWFRNDSSELHESWKYNMFSINVALITINEASAE 7357
QY 95 -----||:|----- 94
Db 7358 DSGDYICEAHNGVGDASCSTALTVRAPPVFTQKPSVGALKGSDVILQCEISGTPPEVV 7417
QY 95 ----- 94
Db 7418 WVKDRQVRNKKFKITTSKHFDTSLHILNLEASDVGEYHCKATNEVSGDTCSCSVKFKEP 7477
QY 95 ----- 94
Db 7478 PRFVKLSDTSLIGDAVELRAIVEGQPIVVVWLKDRGEVIRESENTRISFIDNIATIQ 7537
QY 95 ----- 94
Db 7538 LGSPEASNSGKYICQIKNDAGMRECSAVLTVLEPARIEKPEPMTVTTGNPFALBCVVTG 7597
QY 95 ----- 94
Db 7598 TPESAKWFKDGRSADSADSKHHTFINKVASLKIPCAEMSDKGLYFEVKNVSGKSNCTV 7657
QY 95 ----- 94
Db 7658 SVHVSDRVPPSFIRKLDVNALGASVLECRVSGSAPISVGWFDQGNIEIVSGPKQSS 7717
QY 95 ----- 94
Db 7718 FSENVTNLNLSLEPBDTGIYTCVAANVAGSDECSAVLTQBPSPSQEPTDSVEVLPGMS 7777
QY 95 ----- 94
Db 7778 LFTSVIRTPPFVKWFKGSRVLPVSGESCNISLEDFTVELEFVQPLESGDYSCLVTN 7837
QY 95 ----- 94
Db 7838 DAGSASCTHFLVKBPATPVKRLADFSVETGSPVILEATYTGTPPISVSWIKDEVILISQS 7897
QY 95 ----- 94
Db 7898 ERCSITMTEKSTILELESTIEDYQYSLIENEAGQDICEALVSLPEPPYFIEPLEHVE 7957
QY 95 ----- 94
Db 7958 AVIGEPATLOCKVGTPEIRISWYKEHTKLSAPAYKMQFKNNVASLVINKVDHSDVGEY 8017
QY 95 -----LNKGE----- 99
Db 8018 SCKADNSGAVASSAVLVIKERKLPPFFARKLKVHETLGPVAFECRINGSLEPQVSWY 8077
QY 100 -----:|:|----- 99
Db 8078 KDGVLKDDANLQTSFVHNVAFLQILQTDQSHIGQYNCASNPLGTASSAKLILSEHEV 8137
QY 100 ----- 99
Db 8138 PPFOLKPVSDALGESGTFKCHVTGTAPKITWAKDNREIRPGNGYKMTLVENTATLT 8197
QY 100 ----- 99
Db 8198 VLKVGKGDAGQYCYASNIAKDCSAHLGVQEPFRFIKKLEPSRIVKQDEFTREYCKIG 8257
QY 100 ----- 99
Db 8258 GSPEIKVLWKDETEIQSSKFRMSFVDSVAVLEMHNLSVEDSGDYTCEAHNAAGSASS 8317
QY 100 ----- 99
Db 8318 TSLKVKPEPIFRKKPHPIETLKGADVHLECELOQTTPPHFVSWYKDKRELRSKKYKIMSE 8377

QY 100 ----- 99
Db 8378 NFLTSIHILNVDAADIGEYQCKATNDVGDTCVGSIALKAPPRFVKKLSDISTVTVGKEVQ 8437
QY 100 ----- 99
Db 8438 LQTTIEGAEPISVWVKDKGEIVRESNLIWISYSENIATLQFSRVEPANAGKYTCQIKND 8497
QY 100 ----- 99
Db 8498 AGMQECFATLSVLEPATIVERKESIKVTGTCTTLECTVAGTPELSTKFKDKGKELTSDN 8557
QY 100 ----- 99
Db 8558 KYKISFPNKVSGKLIINVAPSDSGVYSPFVQNPVQKDSCTASLQVSDRTVPSPFTRKLKE 8617
QY 100 ----- 99
Db 8618 TNGLSGSSVWMECKVYGGPPISVSWFHEGNEISSGRKYQTTLTDTNTCALTVNMLEESDSG 8677
QY 100 ----- 99
Db 8678 DYTCTIATNAGSDECSAPLTVREPPSFVQKPDMDVLGTNTFTSIIVKGTPPFSVSWFK 8737
QY 100 ----- 99
Db 8738 GSSELVPGDRCNVSLSDSVALELFDVDTQSGETYCIVSNKASCTTHLYIKAPAKF 8797
QY 100 ----- 99
Db 8798 VKRLNDYSIEKGKPLILEGTFTGTPPISVTWKNGINVTSPORCNITTTKSAILEIPSS 8857
QY 100 ----- 99
Db 8858 TVEDAGQYNCYIENASGKDCSAQILILEPPYFVKQLEPVKSVGSDSASLQCLAGTPEI 8917
QY 100 ----- 99
Db 8918 GVSWMYKGDTKLRPTTYTKMHFNNAVATLVFNQVDINDSGEYICKAENSVEGVSASTFLTV 8977
QY 100 ----- 99
Db 8978 QEQKLPPSFRQLRDVQETVGLPVVFDCAISGSEPSVSWYKDGKPLKDSNVQTSFLDN 9037
QY 100 ----- 99
Db 9038 TATLNIFKTDRLSLAGQYSCYCTATNPIGSASSARLILTEGKNPPFFDIRLAPDAVVGESA 9097
QY 100 ----- 99
Db 9098 DFECHVTCTQPIKYSWAKDSREIRSGGKYQISYLENSAHLTVLKVDKGDGSGQYTCYAVNE 9157
QY 100 ----- 99
Db 9158 VGKDSCTAQLNIKERLIPPSFTKRLSETVEETEGNSFKLEGRVAGSQPITVAVYKNNIEI 9217
QY 100 -----YKCKVSN----- 106
Db 9218 QPISNCEITFKNNTLVQLVRKAGMNDAGLYTCVSNNDAGSALCTSSIVIKEPKKPPVFDQ 9277
QY 107 ----- 106
Db 9278 HLTPTVTSGEYVOLSCHVQSEPIRIQWLKAGREIKPSDRCSFASGTAVLELRVAK 9337
QY 107 ----- 106
Db 9338 ADSDGYVKASNVAGSDTTKSKVTIKDKPAVAPATKAAVDGRLFFVYSEPOSIRVVEKTT 9397
QY 107 ----- 106
Db 9398 ATFTIAKVGDPFIPNVKWTGKWROLNOGGRVFIHQKGDEAKLEIRDTTKTDSGLYRCVAF 9457

```
QY 107 ----- 106
Db 9458 NEHGETESNVNLOVDERKKQIEGDLRAMLKKTPIILKKGAGEEEDIDIMELKNVDPK 9517
QY 107 ----- 106
Db 9518 YEKARYMGITDFRGLQAFELKQSEETHRLEIEIERSERDERKEFEELYSFIQORL 9577
QY 107 ----- 106
Db 9578 SQTEPVLTKIDENQIVLKNDAVFEIDIKINYPEIKLSWKYKTEKLEPSDKFEISIDGD 9637
QY 107 ----- 106
Db 9638 RHLVRKNCOLKQDGNVRLVCGPHIASAKLTVIEPAWERHLQDVTLKEGTCMTQCFSV 9697
QY 107 ----- 106
Db 9698 PNVKSEFRNGRILKPOGRHKTEVEHKVHKLTIADVRAEDQGYTCKYEDLETSALRIE 9757
QY 107 ----- 106
Db 9758 AEPQITKRQIONIVVSEHQSAFECESVDDAIVTWYKGTTELTSQKYNFRNDGRCHYM 9817
QY 107 ----- 106
Db 9818 TIHNVTDDDEGVYSVIARLEPRGEARSTAEYLTLTKLELKPDPIDPSRVPIPTMPIR 9877
QY 107 ----- 106
Db 9878 AVPEEIPPVAPPPIPLLPTPEKKPPKRIEVTKAVKDKAKVVAKPCKEMTPREEIV 9937
QY 107 ----- 106
Db 9938 KKPPPTTLIPAKAPELIDVSSKAEVIMTITRKKEVQKEAEVYKQAVHKEKRVFI 9997
QY 107 ----- 106
Db 9998 ESPEEYDELEVEPYTEPFQPYIEPDEYEEIKVEAKKEVHEEEDPEEQEYERE 10057
QY 107 ----- 106
Db 10058 EGYDEGEEMEAYQREVIVQOKEVYEEESHERRKVPKAVEKKAPPPKVIKKPVIEKIE 10117
QY 107 ----- 106
* Db 10118 KTSRMEEEKVQTKVPEVSKKIVPQKSPTPVOEEVIEVKVPVHTKKMVISSEKMFPA 10177
QY 107 ----- 106
Db 10178 SHTEEEVSVTVPEVQKEIVTEEXIHVAVSKRVEPPPKVPPELPEKPAPEAVPVPKKE 10237
QY 107 ----- 106
Db 10238 PPAPKVPPEKVPKPEKKVPVPVKKPEAPPKVPPEKVPKVPPEKIPVPVAKKKEAPPA 10297
QY 107 ----- 106
Db 10298 KVEVQKRVVTEKITIVTORSESPPAVPEIPKKKVPKEKVPVPRKEEVPPPKVPAL 10357
QY 107 ----- 106
Db 10358 PKKVPPEKVAVPVPAKAPPRAEVSKTVVEERFRVAEEKLSFAVQORVEVTRHEVS 10417
QY 107 ----- 106
Db 10418 AESEWSYSEEGVSVTVREEREEREEAEVTEYEVMEPEPEYVVEEKLHIISKRVAE 10477
QY 107 ----- 113
      : : : : : KALPAPI
Db 10478 PAEVTERQEKKIVLKPIPAKIEEPPPAKVPAPKKTIVPEKKVPAPVPKKEKVPKPVPE 10537
QY 114 ----- 113
```

```
Db 10538 EPKKVPPEKKVPKVIKMEERPLAKVTERHMQITOEKVLVAVTKKEAPPKARVPPEPKR 10597
QY 114 ----- 113
Db 10598 AVPEEKVLKLPKREBEPPAKVTEFRKRVVKEEKVSIAPKRBEPQIKEVTIMEEKRAY 10657
QY 114 ----- 113
Db 10658 TLEEAHSVQRESEYEEYEDYKFEPEYEPTBEYDQYEEYEREYERYSEHEEYITEPE 10717
QY 114 ----- 113
Db 10718 KPIPVKVPPEPVPTKPAKPAKVLKKAPEKVPVPIPKKLKPPPKVPPEPKKVEEK 10777
QY 114 ----- 113
Db 10778 IRISITKREKEQVTEPAAKVPMKPKRVVABEKVPVPRKEVAPVVRVDEVKELEPEEVAF 10837
QY 114 ----- 113
Db 10838 EEEVTHVEEYLYVEEEYIHEEEFITBEEVVPVPIPVKVPVPRKVPPEKKKVPVPKK 10897
QY 114 ----- 113
Db 10898 KEAPPAKVPPEKPKPEKVPVLIPKKEKPPPAKVPPEKVPVPPKVPPEKVPVPAKEAPPA 10957
QY 114 ----- 113
Db 10958 KVPPEKPKVPPEKVPVPAKPAKVEAPPAKVPPEVKLIPBEKKPTVPKPKVEAPPPKVPK 11017
QY 114 ----- 113
Db 11018 KRPVPVPVALPQEEEVLFEEIIVPEEVLPEEVLPEEVLPEEVLPEEVLPEEVLPEEVLPEE 11077
QY 114 ----- 113
Db 11078 EEEVPEEYVPEEVEFVEEVLPEVKPKVPVPAVVPVPEIKKKVTEKKVVPKKEAPPA 11137
QY 114 ----- 113
Db 11138 KVPPEKPKVEEKRIILPKEEVLPEVTEPEPEEPISEBEIPEEPPPSIEVEEVEAPPRVP 11197
QY 114 ----- 113
Db 11198 EVIKKAVPEAPTVPKPKVEAPPAKVSKKIPEEKVPVPVQKKEAPPAKVPPEKVPKVEKKV 11257
QY 114 ----- 113
Db 11258 LVPKKEAVPPAKGRTVLEEKVSVAFROEVVVVKERLELEVVAEVEEIPPEEEFHEVEEYF 11317
QY 114 ----- 113
Db 11318 EGEFHEVEEVEFIKLEQHRVDEEHKRVKHVRVIEVFEAEEVEVEKPAKPKGPEISEKII 11377
QY 114 ----- 113
Db 11378 PPKPPTKVVPKPEPAKVPPEKVIIVVEKVRVPEPRVPPTKVDPDLPPKEVVPKVKV 11437
QY 114 ----- 113
Db 11438 PVPPAKKPEAPPKVPPEAPKVEKVPPEKVPVPPPKKPEVPPTKVPEVPAKAAVPEKKVPEAI 11497
QY 114 ----- 113
Db 11498 PPKPEPPPEVPEAPKVEVPEKKVPAAPPKKDEVTVKVDEAPKVEVVPPEKKVPVPPPKKP 11557
QY 114 ----- 113
Db 11558 EVPTTKVPEVPKVAPEKKVPEAIIPPKPESPPEVEPEEPEEVALEBPPEVVEPEPEAP 11617
QY 114 ----- 113
```

Db 11618 PQVTPPKKVPPEKKAPAVAKKPELPPVKYVEPKVVVPEKKVPLVPPKKPEAPPKVP 11677
QY 114 ----- 113
Db 11678 EYPKEVWPEKKVAVPKKPEVPPAKVPEVPKPKVPLEKPAVPVPERAESPPPEVVEEPEEI 11737
QY 114 -----EK 115
Db 11738 APEEEIAPEEKVPVVAEEEEPEVPPAVPEPKKIIPEKKVPVKKPEAPPKPEPEK 11797
QY 116 TISK----- 119
Db 11798 VIEPKLKRPPPPPPAPPKEDVKEIFQLKAIPKKKVPKQVPEKVELTPLKVPGGEK 11857
QY 120 ----- 119
Db 11858 KVRKLLPERKPEKEEVVLKSVLRPRPEEPEKVPKLEKVKKPAVPEPPPKPVEVE 11917
QY 120 ----- 119
Db 11918 VPTVTKRERKIPEPTKVPEIKPAIPLPAPEPKKPEAEVKTIKPPPVPEPTPIAAPVTV 11977
QY 120 ----- 119
Db 11978 PVBGKAEAKAPKEAAKPKGPIKGVPKKTPSPIEAERRKLRPGSGEKPPDEAPFTYQL 12037
QY 120 ----- 119
Db 12038 KAVPLKFVKEIKDIIITSEFVGSSAIFELVSPSTAITTMKDGNSIRESPKHRFIADG 12097
QY 120 -----AKGP- 124
Db 12098 KDRKLIIDVQLSDAGEYTCVLRGNKEKTSKALVVEELPVRFVKTLEEEVTVVKGQPL 12157
QY 125 ----- 124
Db 12158 YLSCELNKERYVWRKDKIVVEKPGRIVPGVIGLMLRALTIINDADDTAGTYTVTVENAN 12217
QY 125 ----- 124
Db 12218 NLECSKVVEVIRDLVKPIRDQHVPKGTAFACDIADKTPNKKWFKGYDEIPAEPN 12277
QY 125 ----- 124
Db 12278 DKTEILRDGNHLYLKIKNAMPEDIAEYAVEIEGKRYPAKLITIGEREVELLKPIEDVTIYE 12337
QY 125 ----- 124
Db 12338 KESASFDAEISPADIPGQWKLKCELLRPSPTCEIKAEGGKRFLTLRKVKLDQAGEVLYQA 12397
QY 125 ----- 124
Db 12398 LNAITAILTVKEIELFAVPLKDVTVPERQARFECVLTREANVWSKGPDIKSSDKF 12457
QY 125 ----- 124
Db 12458 DIIADGKKHILVINDSQFDEGVYTAIEVEGKTSARLFTVGTIRLAFMSPLEDQTVKEGET 12517
QY 125 ----- 124
Db 12518 ATFYCELSHEKMHVVMFKNDAKLHTSRTVLISSEKTHKLEMKEVTLDDISOIKAQVKEL 12577
QY 125 ----- 124
Db 12578 SSTAQLKVEADPYFTVKLHDKTAVEKDEITLKEVSKDVPVKWFKDGEIIVPSPKYSIK 12637
QY 125 ----- 124
Db 12638 ADGLRILKIKKADLKDKGEYVDCGTDKTRANVTVEARLIKVEKPLYGEVFGETAHF 12697
QY 125 ----- 124
Db 12698 EIELSEPDVHGQWKLKGQPLTASPDCEIIEDGKKHILHNCQLGTMTEVGSFQAANAKSA 12757

QY 125 -----REPO----- 128
Db 12758 ANLKVKELPLIFITPLSDVKVFEKDEAKFECEVSREPXTFRWLKGTQIEITGDDREFELIKD 12817
QY 129 ----- 128
Db 12818 GTKHSWIKSAAFEDEAKYMPAEDKHTSGKLIIEGIRLKFUTPLKDVTAKEKESAVFTV 12877
QY 129 ----- 128
Db 12878 ELSHDNIIRVKFKNDORLHTRTSVSMQDEGKTHSTFKDLSIDDTSQIRVEAMGMSSEAK 12937
QY 129 ----- 128
Db 12938 LTVLEGDYPFTGKLQDYTGVEKDEVILQCEISKADAPVKWFKDGKEIKPSKNAVIAKDGK 12997
QY 129 ----- 128
Db 12998 KRMLILKKALKSDIGQYTCDCGTDKTSGKLDIEDREIKLVRPLHSVEVMEETARFETEI 13057
QY 129 ----- 128
Db 13058 SEDDIHANWKLKGEALLQTPDCEIKEGKIHSVLHNCRLDQGTGVDFQAAVNVKSAHLR 13117
QY 129 ----- 128
Db 13118 VKPRVIGLLRPLKDVTVTAGETATFDCELSYEDIPVENYLGKKKLEPDKVVPVPRSEKGVH 13177
QY 129 ----- 128
Db 13178 TLTLRDVKLEADAGEVQLTAKDFKTHANLFVKPEPPVEFTKPLEDQTVBEGATAVLECEVSR 13237
QY 129 ----- 128
Db 13238 ENAKVWFKNGTEILSKKYEIVADGRVKLVIHDPEDIKTYTCDAKFKTSCNLNV 13297
QY 129 ----- 128
Db 13298 PPVFEFLRPLDLQVREKEMARFECELSRENKAKWFKDGAEIKGKKYDIISKAVRIL 13357
QY 129 ----- 128
Db 13358 VINKCLLDEAEYSCEVRTARTSGMLTVLEEAFTKNLANIEVSETDTIKLVCBVKPG 13417
QY 129 ----- 128
Db 13418 AEVIWYKGDEEIIETGRVEIILTEGRKRILVIONAHLEDAGNVNCRLPSSRTDGVKVKVHEL 13477
QY 129 ----- 128
Db 13478 AAEFISKPONLEILEGEKAEFVCSISKESFPVOWKRDDKTLESGBKYDVIADGKKRVLV 13537
QY 129 ----- 128
Db 13538 KDATLQDMGYVMVGAARAAHLTVIEKLRIVPLKDRVKEQOEVEVFNCEVNTGAKA 13597
QY 129 -----VYTL----- 132
Db 13598 KWRNEEAIKFDSSKYIILQKDLVYTLRIRDAHLDDOANVNSLTNHRGENVKSAANLIVE 13657
QY 133 ----- 132
Db 13658 EEDLRIVEPLKDIEIEMKKSVTFWCKVNRNLNVTLWTKNGEVPFDNRVSRVVDKYKIHL 13717
QY 133 ----- 132
Db 13718 TIKDCGFPDEGEYIVTAGODKSAVELLIIEAPTEFEVHELDQTVTEFDQAVFSCOLSREK 13777
QY 133 ----- 132
Db 13778 ANVKWYRNGREIKEGKKYKFEKDGSIHRLIITKDRLDDECEYACCGVEDRKSARLFEVEI 13837

QY 133 ----- 132
Db 13838 PVEIIRPPQDILEAPGADVFLAELNKKDKEVEQWLRNNMVVVOGDKHOMSEKIHRLQI 13897
QY 133 ----- 132
Db 13898 CDIKPRDQGEYRFIAKDEARAKELAAAPKIKTADQDLVVDVGKPLTMVVPYDAYPKAE 13957
QY 133 ----- 132
Db 13958 AEWFKENEPLSTKIDTTAEQTSFRILEAKKGDKGRYKIVLQNKHGAEGFINLVIDVP 14017
QY 133 ----- 132
Db 14018 GPNRLEVTETFDGEVSLAWEEPITDGGSKIIGVVVERRDIKRKTWVLATDRAESCFTV 14077
QY 133 ----- 132
Db 14078 TGLQKGGVEYLFRVSAARNVGTGEPVETDNPVEARSKYDVGPPNLNVTITDVNRFGVSLT 14137
QY 133 ----- 132
Db 14138 WEPPEYDGAETNYVIELRDKTSIRWDTAMTVRAEDLSATVTDVVEGQEYSFRVRAQNR 14197
QY 133 ----- 132
Db 14198 IGVGKPSAATPFVKVADPIERPSPVNLTSDDQTQSSVLKWEPLPKDGGSPILGYIER 14257
QY 133 ----- 132
Db 14258 CEEGKNWIRCNMKLVPETYKVTGLEKGNKYLYRVAENKAGVSDPSEILGPLTADDAF 14317
QY 133 ----- 132
Db 14318 VEPTMDLSAPKDGLEIVPNPITILVPSTGYPRPTATWCFGDKVLETGDRVKMKTL SAYA 14377
QY 133 ----- 132
Db 14378 ELVISPERSDKGIYTLKLENRVKTSIGEIDVNVNVIARPSAPKELKFGDITKDSVHLTWP 14437
QY 133 ----- 132
Db 14438 PDDGGSPLTGYVVEKREVSRTWTVMDFVTDLEFTVDPDLVQGEYLFKVCARNKCGPG 14497
QY 133 ----- 132
Db 14498 EPAIVDEPNMSTPATVPDPENVKWRDRTANSIFLTWDPKNDGGSRIKGYIVERCPRG 14557
QY 133 ----- 132
Db 14558 SDKWVACGEPVAETKMEVTGLEGKWYAYRVKALNRQASKPSRPTETIQAVDTQEAPEI 14617
QY 133 ----- 132
Db 14618 FLDVKLLAGLTVKAGTKIELPATVGTGPEPKITWTKADMILKODKRITIEVNPVKSTVTI 14677
QY 133 ----- 132
Db 14678 VDSKRSDTGYIIIEAVNVCGRATAVEVNVLDKPGPAAFDITDVTNESCLLTWNPPRDD 14737
QY 133 ----- 132
Db 14738 GGSKITNVVERRATDSEVNHKLSSTVKDTNFKATKILIPNKEYIFRVAANMNYGVGEPVQ 14797
QY 133 ----- 147
|||: :|||
---PPSR---DELTKNOVSLT
Db 14798 ASPITAKYQDPGPGPTLEPSPDITKDAVTLTWCEPDDGGSPITGYWVERLDPDTDKW 14857
QY 148 ----- 147
Db 14858 RCNMPVKDITYRVKGLTNKKYFRVLAENLAGPGPKSSTEPILLIKDPIDPWPFGPKP 14917
QY 148 ----- 147

Db 14918 TVKDVGKTSVRLNWKTRPEHDGAKIESYVIELMLKTGTDEWVRVAEGVPTTQHLLPGLMEG 14977
QY 148 ----- 147
Db 14978 QEYSFRVRAVNKAGESEPSDFVLCREKLYPPSPRWRLEVINITKNTADLKWVPEKD 15037
QY 148 ----- 147
Db 15038 GGSPITNYIVEKRDVRKGMQTVDTTVKDTKCTVTPLTGSLYVFRVAENAIGQSDYTE 15097
QY 148 ----- 147
Db 15098 IEDSVLAKDTFTTGPFPYALAVDVTKRHVDLKWEPPKNDGGRPIQRYVIEKKERLGRW 15157
QY 148 ----- 147
Db 15158 VKAGTAGPCNFRVTDVIEGTEVOFVRAENAGVGHPSPTIEILSIEDPTSPSPPLD 15217
QY 148 ----- 147
Db 15218 LHVTDAGRKHIAIAWKPPKNGGSPILGYHVEMCPVGTCKMWRVNSRP IKDLKFKVEGV 15277
QY 148 ----- 147
Db 15278 VPKYVYLRVAVNAIGVSEPSSEISENVAKDPCKPTIDLETHDIIIVIEGKLSIPVPF 15337
QY 148 ----- 147
Db 15338 RAVPVPTVSHKDGKEVKASDRITMKNDHISAHLEVPKSVRADAGIYITILENKLGSATA 15397
QY 148 ----- 147
Db 15398 SINVKVIGLPCKDIKASDITKSSCKLTWEPPEFDGPTILHYVLERREAGRRTYIPVM 15457
QY 148 ----- 147
Db 15458 SGENKLSWTVDLIPNGEYFFRVKAVNKVGGGEYIELKNPVAQDPKQPPVPDVEVHN 15517
QY 148 ----- 147
Db 15518 PTAEMTITWKPPPLYDGGSKIMGYIEKIAKGEERWRCNEHLVPILTYTAKGLEEGKEY 15577
QY 148 ----- 147
Db 15578 QFRVRAENAAAGISEPSRATPPTKAVDPIDAPKVILRTSLEVKRGDEIADASISGSPYPT 15637
QY 148 ----- 147
Db 15638 ITWIKDENVIPEEIKKRAAPLVRRRKGEVQEEFPVLPLTQRLSIDNSKKGESQLRVRD 15697
QY 148 ----- 147
Db 15698 SLRPDHGLYMIKVENDHGIAKAPCTVSVDTPGPPINFPVEDIRKTSVLCKWEPLDDGG 15757
QY 148 ----- 147
Db 15758 SEIINYTLEKKDKTKPDSEWIVVTSTLRHCKYSVTKLIEGKEYLFRVRAENRFGPGPCV 15817
QY 148 ----- 147
Db 15818 SKPLVAKDPFGPPADPKPIVEDVTSNMLYKWNPKDNGSPILGYWLEKREVNSTHWSR 15877
QY 148 ----- 147
Db 15878 VNKSLNALKANVDGLLEGLTYVFRVCAENAAAGPKFSPSPDKPTAHDPISPPGPIPRV 15937
QY 148 ----- 147
Db 15938 TDTSSTTIELEWEPAPFNGGGEIVGYFVDKQLVGTNEWSRCEKMIKVRQYTVKEIREGA 15997
QY 148 ----- 147

Db 15998 DYKLRVSAVNAGEGPGTQPTVVAEPQEPAPVELDVSKGGIQIMAGKTLRIPAVVTG 16057
QY 148 ----- 147
Db 16058 RPVPTKVTKEGELDKDRVVIDNVGTKSELIIDKALRKHGRYVITATNCSGSKFAAAR 16117
QY 148 ----- 147
Db 16118 VEVFDVPGVLDLKPVTNRKMLNWSODPEDDGSSEITGFIERKDAKMHWTWRQPIETE 16177
QY 148 ----- 147
Db 16178 RSKCDITGLLEGOEYKFRVIANKFGCGPPVEIGPILAVDPLGPTSPERLTYTERTKST 16237
QY 148 ----- 147
Db 16238 ITLDWKEPRNGSGSPIQGYIIIEKRRHDKDPFRVNRRLCTPTTSFLVENLDEHOMYFRVK 16297
QY 148 ----- 147
Db 16298 AVNEIGESEPSLPNVVQDDEVPTIKLRLSVRGDTIKVKAGEPVHIIPADVTGLPMPKI 16357
QY 148 ----- 147
Db 16358 EWSKNETVIEKPTDALQITKEEVSRSSEAKTELSIPKAVREDKGTVTYTASNRLGSVFRNW 16417
QY 148 ----- 147
Db 16418 HVEVDRPSPRNLAVTDIAKESCYLTWDAPLDNGGSETHYVIDKRDASRKAWEVET 16477
QY 148 ----- 147
Db 16478 NTAVERKRYGIWKLIPNGQYEFVRVAVNKYGISDECKSKVVIQDPYRLPGPGPKVLAR 16537
QY 148 ----- 147
Db 16538 TKGMLVSWTPPLDNGGSPITGYWLEKREEGSPYWSRVSRAPIITKVGLKGVEFNVPRLLLE 16597
QY 148 ----- 147
Db 16598 GVKYQFRAMAINAAGIGPSESDPEVAGDPIFPFGPPSCPEVKOKTKSSISLGNWPKPAK 16657
QY 148 ----- 147
Db 16658 DGGSPIKGYIVEMQEBGTTDKRVNRPDKLITTCWCVPNLKELRYFRVRAVNEAGES 16717
QY 148 -----CLV----- 150
Db 16718 EPSDITGEIPANDIQEEPEVFIIDIGAQCCLVCKAGSQIRIPAVIKGRPTPKSSWEFDGKA 16777
QY 151 ----- 150
Db 16778 KRAMDGVHDIPEDAQLETAENSSVIIIECKRSHTKGYSITAKNKAGQKTANCVRKVM 16837
QY 151 -----KGF----- 153
Db 16838 VPGPKDLKVSIDITRSGRLSNKMPDDGGDRIKGVIEKRTIDGKAWTKVNPDCGSTTF 16897
QY 154 -----YPSD----- 157
Db 16898 VVPDLLSEQOYFRVRAENRFGIGPPVETIQRTTARDPIYPPDPPIKILIGLITKNTVHL 16957
QY 158 ----- 157
Db 16958 SMKPPKNDGSPVTHVIVECLAWDTGTKEAWROCNRKRDVEELQFTVEDLVEGGEYEP 17017
QY 158 ----- 157
Db 17018 VKAVNAAGVSKPSATVGPCDCQRPMPSPSIDLKEEFMEVEBGTNVNIVAKIKGVFPPTLTW 17077
QY 158 ----- 157
Db 17078 FKAPPKPDNKEPVLVDTHVNLVDDTCLTIVIPOSRRSDTGLTYTITAVNNIGTASKEMR 17137

QY 158 ----- 157
Db 17138 LNVLGRPGPPGPIKPFESVSADOMTLSNFPKDDGGSKITNVVIEKREANRKTWVHSSE 17197
QY 158 ----- 157
Db 17198 PKECTYTI PKLLEGEHYVFRIMAQNKYGIGELDSEPTARNLFSVPGAPDKPTVSSVTR 17257
QY 158 --IAVEWE----- 163
Db 17258 NSMTVNWEEPEYDGGSPVTGYWLEMKDTSKRWKRVNRDPIKAMTILGVSYKVTGLIEGSD 17317
QY 164 ----- 163
Db 17318 YQFRVYAINAAGVGPASPLSDPATARDPIAPGPPPKVTDWTKSSADLEWSPPLKDGGS 17377
QY 164 ----- 163
Db 17378 KVTGYIIVEYKEBGEKEWEKGDKEVRGTKLVVTGLKEGAFYKFRVSAVNIAGIGEPGEVT 17437
QY 164 ----- 163
Db 17438 DVIEMKRLVSPDLQLDASVRDRIVVHAGGVIRIIAYVSGKPPPTVTWNMNERLTLPQEA 17497
QY 164 ----- 163
Db 17498 IETTAISSMWIKNCORSHQGVYSLAKNEAGERKKTIIIVDVLDPGPPVGTPTFLAHLNLTN 17557
QY 164 -----SNGOPENNY----- 172
Db 17558 ESCKLTWESPEDDGGSPITNYVIEKRESRRRAWTPTYTYVTTRQNATVQGLIOGKAYFFRI 17617
QY 173 ----- 172
Db 17618 AAENSIGMGPFVETSEALVIREPITVPERPEDLEVKEVTNKTVTLTWNPYPKYGGESEIN 17677
QY 173 ----- 172
Db 17678 YVLESRLIGTEFKHTNDNLLSRKYTVKGLKEGDTYEVRSVAVNIVGOGKPSFCTKPTT 17737
QY 173 -----KTT-----KTT 175
Db 17738 CKDELAPPTLHLDFRDKLTIRVGEAFALTGRYSGKPKVSWFKDEADVLEDDRTHIKTT 17797
QY 176 P----- 176
Db 17798 PATLALEXIKAKRSDSGKYCVVVENSTGSRKGCQVNVVDRPGPPVGVPSFDEVTKDMV 17857
QY 177 ----- 176
Db 17858 ISWKPPLDGGSKITNYIIIEKKEVGKDVMPVTSASAKTCKVSKLLEGKDYIFRIHAEN 17917
QY 177 ----- 176
Db 17918 LYGISDPLVSDSMKAKDRFRVPDAPDQPIVTEVTKDSALVTWNKPHDGGKPTITNVILEKR 17977
QY 177 ----- 176
Db 17978 ETMSKRWARVTKDPIHPYTKFRVPDLLEGQCYEYFRVSAENEIGIDPSPSPKPVFAKOPI 18037
QY 177 ----- 176
Db 18038 AKPSPVNPPEALDITTCNSVDLTWQPPRHGGSKILGYIVEYQKVGDEEWRANHTPESCP 18097
QY 177 ----- 176
Db 18098 ETKYKVTGLRDCQTYKFRVLAVNAAGESDPAHVPEPVLVKDRLEPPELLIDANMAREQHI 18157
QY 177 ----- 176
Db 18158 KVGDTLRLSAIIKGVFPFKVTKWKKEDRDAPTAKRIDVTPVGSKLEIRNAHEDGGIYSUT 18217

```
QY 177 ----- 176
Db 18218 VENPAGSKTVSVKVLVDKPGPPDLLEVSEIRKDCSVLTWKEPLDDGGSVITNVVVERRD 18277
QY 177 ----- 176
Db 18278 VASAQMSPLSATSKKKSHFAKHLNEGNOYLFRVAENQYGRGPFVETPKPIKALDPLHPP 18337
QY 177 ----- 176
Db 18338 GPPKDLHHVDVKTEVSLVWKNKPDROGGSPITGYLVEYQEEGTQDWIKFKFTVTNLECVVT 18397
QY 177 ----- 176
Db 18398 GLOQGYFRFRVAENIVGLGLPDTTPIECQEKLVPPSVELDVKLTIEGLVWKAGTTVRF 18457
QY 177 ----- 176
Db 18458 PAIRGVVPFTAKWTDGSEIKTDEHYTVETDNFSSVLTITKNCLRRDTGTYQITVSNAG 18517
QY 177 ----- 176
Db 18518 SKTVAVHLTVLDVPGPTGPIINILDTPEHMTISWQPPKDDGGSPVINIYIVEKQDTRKDT 18577
QY 177 ----- 176
Db 18578 WGVSSGSSKTKLIPHLOKGCYVFRVAENKIGVPPDLDSTPTVAKHKFSPPPGKP 18637
QY 177 ----- 176
Db 18638 VVTDITENAAVTWTLPKDGGSPITGYMERREVTGKWVRNKTPIADLKFRVTGLYEG 18697
QY 177 ----- 176
Db 18698 NYEYFRVAENLAGLSKPSDDPIKACRPIKPPGPDINPKLKDKSRETADLVWTKPLSD 18757
QY 177 ----- 176
Db 18758 GGSFILGYVVECQKPGTAOWNRINKDELIRQCAFVPLGIEGNEYRFRIKAANIVGEGEP 18817
QY 177 ----- 176
Db 18818 RELAESVIAKDILHPPPEVELDVTCDVITVRVGQTIIRILARVGRPEPDITWTKGVLY 18877
QY 177 ----- 176
Db 18878 REKRVDLIQDLPRVELOIKEAVRADHGKYIISAKNSSGHAQGSAINVNLDRPGPCQNLKV 18937
QY 177 ----- 176
Db 18938 TNVTKENCTISWENPLDNGGSEITNFIVEYRKPQKQWSIVASDVTKRLIKANLANNEY 18997
QY 177 ----- 176
Db 18998 YFRVCAENKVGOGPTIETKPTILAINPIDRPGEPENLHIAADKGKTFVYLKWRRPDYDGS 19057
QY 177 ----- 176
Db 19058 PNL5YHVERBKSGDDWHRVHKSIGKETHYMWDRVCVENQIYEFVRQTKNEGGSOWVYTE 19117
QY 177 ----- 180
Db 19118 EVVVKEDLPVLDLKL5GLTVKAGDTIRLEAGVRGKPPPEVAWTKDKDATDLTRSPTV 19177
QY 181 -----SD-----GSFFLYSKLTV----- 193
||| ||| | : |
Db 19178 KIDTRADSSKFSLTAKRSDGGKYVVTATNTAGSFAYATVNLDPGPVRNLKIVDVSS 19237
QY 194 ----- 193
Db 19238 DRCTVCWDDPDDGGCEIQNYILEKCTKRMWSTYSATVLTGCTVTRLIEGNEYIFRV 19297
QY 194 -----DK----- 195
```

```
Db 19298 RAENKIGTGPPTESKPVIAKTKYDKPGRPDPPEVTKVSKBEETVVMNPPBYDGGKSITGY 19357
QY 196 ----- 195
Db 19358 FLEKKEKHSTRWVPVVKSAIPERRMKVONLLPDHEYOFRVKAENEIGIGEPSLPSRPVVA 19417
QY 196 ----- 195
Db 19418 KDPIEPPGPTNFRVVDTTKHSITLGMGKPVYDGGAPLIGYVVMRPFKIADASPDEGWR 19477
QY 196 ----- 195
Db 19478 CNAQAQLVRKEFTVTSIDENQYEFVRCAQNOVGIGRPAELKAIPKEILEPEIDLDA 19537
QY 196 ----- 195
Db 19538 SMRKLIVRACCPTRLFAIVRGPRAPKVTRKVGIDNVVRKGQVDLVDTHAFLVIPNSTR 19597
QY 196 ----- 195
Db 19598 DDSGKSLTLVNPAGEKAFVNVVRVLDTPGPVSDLKVSDVTKTSCHVSWAPPENDGGSQV 19657
QY 196 ----- 195
Db 19658 THYIVKREADRKTWSTVTPEVKKTSFHTNLPVGEYIFRVTAVNEYGPGVPTDVPKPV 19717
QY 196 ----- 195
Db 19718 LASDPLSEDPBRKLEVTEMTKNSATLAWLPLRDGGAKIDGYITSYREEBPADRWTY 19777
QY 196 ----- 195
Db 19778 SVVKDLSLVTLGLKEGKKYKFRVAARNAGVSLPREAEGVVEAKEQLLPKILMEQITI 19837
QY 196 -----SRWQOQN----- 202
||| : |||
Db 19838 KAGKLRIEAHVYKGPHTCKWKKEDEVVTSSHLAVHKADSSILIIKDTRKDSGYYS 19897
QY 203 ----- 202
Db 19998 LTAENSSGTDQTKIVVMDAPGPPQPFDISDIDADACSLSWHIPLDGGSNITNTIVE 19957
QY 203 ----- 202
Db 19958 KCDVSRGDWTALASVTKTSRVGKLLIPGQYIFRVRARENFGISEPLTSPKVAQPFEG 20017
QY 203 ----- 202
Db 20018 VPSEPKNARVTKNKDCIFVANDRPDSGGSPILIGYLIERKERNSLWVKANDTLVRSTE 20077
QY 203 ----- 202
Db 20078 YPCAGLVEGLEYSPIRIYALNKGSSPKPTEYVTARMPVDPDPGKPEVIDVTKSTVSLIW 20137
QY 203 ----- 202
Db 20138 ARPKHDGSKIIGYFVEACKLPQDKWVRCNTAPHQIQEEYATATGLEEKAQYQFRAIART 20197
QY 203 ----- 202
Db 20198 AVNISPPSEPSDPVTILAENVPPRIDLSVAMKSLITVKAGTNVCLDATVFGKMPPTVSWK 20257
QY 203 ----- 202
Db 20258 KDGTLLKPAEIGIKMAMORNLCITLFSVNRKDSGDTYITAENSSGSKSATIKLKVLDKPG 20317
QY 203 -----VFSGSV- 208
||| : |||
Db 20318 PPASVKINKMYSDRAMLSWEPPELDDGGSEITNYIVDKRETSRPNWAQVSATVPITSCSV 20377
QY 209 ----- 208
```

Db 20378 KLEGEHYQFICAENKYGVDVETEPAPAIKPNYPDPGRCDPPVISNITKDHMTVSMKP 20437
QY 209 ----- 208
Db 20438 PADDGSPITGYLLEKRETOAVNWKVNKRKPIERTLKATGLOEGTEYEFRTAINKAGP 20497
QY 209 ----- 208
Db 20498 GKPSDASKAAYARDQYPPAPPAFPKVVYDTRSSVSLSWGKPAYDGGSPIIIGLYVEVKRA 20557
QY 209 ----- 208
Db 20558 DSDNWWRCNLONLQKREFTVGLMEDTOYQFRVAVANKIGYSDPSDVPDKHYPKDILIP 20617
QY 209 ----- 208
Db 20618 PEGELDADLRKTLILRAGVTMRLYVPVKGRRPPPKITWSKPNVNLDRIGLDIKSTDEDTF 20677
QY 209 ----- 208
Db 20678 LRCENVNKYDACKYILTLSENSCGKKEYTIVVKVLDTPGPPVNVTVKEISKOSAYVTWEP 20737
QY 209 ----- 208
Db 20738 IIDGSPIIINYVQKRDARKSWSTVTTECSKTSFRVANLEBKSYFFRFAENEYIGID 20797
QY 209 ----- 208
Db 20798 PGETRDANKASQTPGVVDLKVRSVSKSSCSIGWKKPHSDGSGRIIGYVVDFTENKWO 20857
QY 209 ----- 208
Db 20858 RVKMSLSQYSAKDLTEGKEYTFRVSAENENGEGTSEITVVARDVVPDLKGLPDL 20917
QY 209 ----- 208
Db 20918 CYLAKENSNFLUKIPKGPAPSVSWKKGEDPLATDTRVSVESSAVNTTLIVYDCKSDA 20977
QY 209 ----- 208
Db 20978 GXYTITLKNVAGTKEGTISIKVVGKPGIPTGPIKFEVTAEAMTLKWAPPKDDGSEITN 21037
QY 209 ----- 214
MHEALH
: : :
Db 21038 YILEKRDVNNKWNVCASAVQKTFTRVTRLHEGMEYTFRVSANKEYGVGEGKSEPIVAR 21097
QY 215 ----- 214
NHY
: : :
Db 21098 HPFDVPDAPPPNIVDVRHDSVSLTWTDPKKTGGSPITGYHLEFKERNSLWKRANKTPI 21157
QY 215 ----- 214
Db 21158 RMRDFKVTCLTBSGLEVEFRVMAINLAGVKPSLPSEPVVVALDPIDPPGKPEVINITNSV 21217
QY 215 ----- 214
Db 21218 TLIMTEPKYDGGHKLTYIVEKRDLPKSWMKANHVNVPECAFTVTDLVEGKYEFIRA 21277
QY 215 ----- 214
Db 21278 KNTAGAISAPSESTETIICKDEYEAPTIVLPTIKDGLTIKAGDTIVLNAISILUGKPLPK 21337
QY 215 ----- 214
Db 21338 SWSKAGDIRPSDITQITSTPSSMLTIKYATRKDAGEYITATNPFGTKVEHKVTVL 21397
QY 215 ----- 214
Db 21398 DVPGPPGVEISNVAEKATLTWTPLEDGGSPIKSYILEKRETSRLLTWTVVSEDIQSCR 21457
QY 215 ----- 217
NHY
: : :
Db 21458 HVATKLIQOGNEYIFRVSANVHYGGEVPQSEPVKMDVRFGPPGPEKPEVSNVTKNATV 21517

QY 218 ----- 217
Db 21518 SWKRPVDDGGSEITGYHVERREKKSLRVVRAIKTPVSDLRCKAVTGLQEGSTYEFVRSAEN 21577
QY 218 ----- 225
-TKSLSLS-
: : :
Db 21578 RAGIGPPEASDSVLMKDAAYPPGPPSNPHVTDTTKKSASLAWKPHYDGGLEITGYVVE 21637
QY 226 ----- 225
Db 21638 HOKVGDEAWIKDTGTALRITQFVVPDLQTKYKYNFRISAINDAGVGEPAVIPDVEIVER 21697
QY 226 ----- 225
Db 21698 EMAPDFELDAELRRTLIVVRAGLSIRIFVPIKGRPAPEVTWTKDNINLKNRANIENTESFT 21757
QY 226 ----- 225
Db 21758 LLIIEPCNRYDTGKFVMTIENPAGKSGFVNVRLDTPGPVNLNRPDITDKDSVTLHMDL 21817
QY 226 ----- 225
Db 21818 PLIDGSRITNIVEKREATRKSYSSTATTKCHKCTYKVTGLSEGCYFFRVMANENYIG 21877
QY 226 ----- 225
Db 21878 EPTETTEPVKASEAPSPDLSNIMDITKSTVSLANPKPKHDGSKITGVVIEAQRKGSQ 21937
QY 226 ----- 225
Db 21938 WTHITTVKLECVVRNLTEGEYTFQVMVNSAGRSAPRESRPVIVKEQTMPELDLRCI 21997
QY 226 ----- 225
Db 21998 YOKLVIKAGDNIKVEIPVLGRPKPTVTKKGQDILKQTRVNFETTATSTILNINECVR 22057
QY 226 ----- 225
Db 22058 SDSGYPPLTARNIVEGVDVITIQVHDIPGPTGPIKFEVSSDFVTSWDPPEGGVP 22117
QY 226 ----- 225
Db 22118 ISNVVEMRQDSTTWELATTVIRTTYKATRLTTGLEVQFRVKAQNRVGVGITSACI 22177
QY 226 ----- 225
Db 22178 VANYPFKVPDPPGTPQVTAVTKDSMTISWHEPLSDGSGPILGYHVERKERNGILMQTVSK 22237
QY 226 ----- 225
Db 22238 ALVPGNIFKSSGLTDGIAVEFRVIAENMAGKSKPSKPSEPMALDPIPPGKPVPLNITR 22297
QY 226 ----- 225
Db 22298 HTVTLKWAKPEYTGFKITSYIVEKRDLPNGRWLKANFSNILENEFTVSGLTEDAAEYFR 22357
QY 226 ----- 225
Db 22358 VIAKNAAGAISSPPSESDAITCRDDVEAPKIKVDVKFDVTILKAGEAFRLEADVSGRRP 22417
QY 226 ----- 225
Db 22418 PTMEWSKDGKELEGTAKLEIKIADPSTNLVNKDSRRDSGAYTLTATNPGGFAKIFNVK 22477
QY 226 ----- 225
Db 22478 VLDPRGPPPEGLAVTEVTSKCVLSWFPPLDDGGAKIDHYIVQKRETSRLAWTNVASEVQ 22537
QY 226 ----- 225
Db 22538 VTKLKVTLLKNGNEYIFRVMVANKYGVGEPELSEPVLAENVYPPDPKPNBEVTTITKDS 22597

QY 226 ----- 225
Db 22598 MVMGHPDSDGSEIINYIVERRDKAGQWIKCNKKTLDRLRYKVSGLTEGHEYEPRI 23657
QY 226 ----- 225
Db 22658 AENAGISAPSPSPFYKACDVFVKPGPPGNPRVLDFSRSSISIAWKNPIYDGGSEITGY 22717
QY 226 ----- 225
Db 22718 MVEIALPEDEWQIVTPPAGLKATSYITIGLTENQEKIRIYAMNSEGLGEPALVPGTPK 22777
QY 226 ----- 225
Db 22778 AEDRMLPETELDADLRKVVITIRACCTLRLFPVKRPAPEVKWARDHGESLDKATEST 22837
QY 226 ----- 225
Db 22838 SSVTLIVGNVRPDSKYILTVENSSGSKSAFVNRVLDTPGPDOLKVEVTKTSVTL 22897
QY 226 ----- 225
Db 22898 TWDPPLLDGGSKIKNYIVEKRESTRKAYSTVATNCHTSMKWVDOLOQEGCSYYFRVLAENE 22957
QY 226 ----- 228
Db 22958 YGIGLPAETAESKASERPLPPGKITLMDVTRNSVLSWEKPEHDGSRILGIYIVEMQTK 23017
QY 229 ----- 228
Db 23018 GSKWATCATVKVTEATITGLIOGEYSFRVSAQNEKGISDPROLSVPVIAKDLVIPPAP 23077
QY 229 ----- 228
Db 23078 KLLFNTFTVLAGEDLKVDVDFPIGRPTPAVTHKDNVPLKQTRVNAESTENNSLLTTIKDA 23137
QY 229 ----- 228
Db 23138 CREDVGHVVKLNSAGEAETELNVILDKPGPTGPVKMDEVTAADITLSWGPVKYDGG 23197
QY 229 ----- 228
Db 23198 SSINNYIVEKRDSTTTWQIVSATVARTTIKACRLKTGCEYQFRIAAENRYGKSTYLNSE 23257
QY 229 ----- 228
Db 23258 PTVAQYFVKVPGPPGTPVVTLSRDSMEVQWNEPISDGGSRVIGYHLERKERNILWKL 23317
QY 229 ----- 228
Db 23318 NKTPQTKFKTGLEGEVEFRVSAENIVGIGKPSKSECVYARDPCDPPGRPEALIV 23377
QY 229 ----- 228
Db 23378 TRNSVTLOWKKPYDGGSKITGYIVEKKELPEGRWMAKSFNTIIDTFEVTGLVEDHRYE 23437
QY 229 ----- 228
Db 23438 FRVIARNAAGVFSEPESTGALTARDEVDPPIRISMDPKYKDTIVVHAGESFKVDADIYK 23497
QY 229 ----- 228
Db 23498 PIPTIOWIKGDELNTARLEIKSTDFATSLSKVDVAVRDSGNVILAKKNVAGERSVTVN 23557
QY 229 ----- 228
Db 23558 VKVLDREPPEGPVVISGVTAECTLAWKPPLODGGSDIINYIVERRETSRLVWTVVDAN 23617
QY 229 ----- 228
Db 23618 VQTLSCVKTKLLEGNEYTFRIMAVNKYGVGEPESEPVVAKNPVVPDAPKAPVTTVTK 23677
QY 229 ----- 228

Db 23678 DSMIVWERPASDGGSEILGYVLEKRDKEGIRWTRCHKRLIGELRLRVTGLIENHDYEFR 23737
QY 229 ----- 236
Db 23738 VSAENAGLSEPPSPSAYQKACDPIYKPPNNPKVIDITRSSVFLSWSKPIYDGGCEIQ 23797
QY 237 ----- 236
Db 23798 GYIVEKCDVSGEWTMCTPTGTINKNTIEVEKLEKHEYNEFRICAINKAGVGEHADVPGP 23857
QY 237 ----- 236
Db 23858 IIVEEKLEAPDIDLDELRLKIINIRAGGSLRFLVPIKGRPTPEVKWGVDEIRDAALID 23917
QY 237 ----- 236
Db 23918 VTSFTSLVLDNVNRYDSGKYTLFLENSSGTKSAFVTVRVLDTSPPVNLKVITITKDSV 23977
QY 237 ----- 236
Db 23978 SITWEPPLDGGSKIKNYIVEKREATRKSAAVVTNCHKNKSWKIDLOEGCSYYFRVTA 24037
QY 237 ----- 240
Db 24038 NEYGIGLPAQTADPIKVAEVPQPPGKITVDVTRNSVLSWTKPEHDGGSKIIQIVIVEMQ 24097
QY 241 ----- 240
Db 24098 AKHSEKSECARVKSLOAVITNLQTGBEYLFRRVVAVNEKGRSDPRSLAVPIVAKDLVIEP 24157
QY 241 ----- 240
Db 24158 DVKPAFSSYVQVQCDLKIEVPISGRPKPTITWTGDLPLKQTRINVTDSLDTLTSIK 24217
QY 241 ----- 240
Db 24218 ETHKDDGGYGITVANVVQKTAIEIWTLDKPPKPKVFDVSAESITLSWNPPLYT 24277
QY 241 ----- 240
Db 24278 GGCQITNYIVOKRDTTWTVDWSATVARTTLKVTKLKTGTEYQFRIFAENRYGOSFALE 24337
QY 241 ----- 240
Db 24338 SDPIVAQYPYKEPGPPGTPFATAISKDSMWIQWHEPVNNGGSPVIGYHLERKERNILWT 24397
QY 241 ----- 240
Db 24398 KVNKTIIDHTQFKAQNLLEGIEYEFRRVAENIVGVGKASKSECVYARDPCDPPGTPEPI 24457
QY 241 ----- 240
Db 24458 MVRKNEITLOWTRKPYDGGSMITGYIVEKRDLPDGRWMAKSFNTVITQFTVSGLTEDQR 24517
QY 241 ----- 240
Db 24518 YEFVRIAKNAAGATSKPSDSTGPITAKDEVELPRISMDPKFRDITVYNAGETFRLEADVH 24577
QY 241 ----- 240
Db 24578 GKPLPTIWLKRGDKEIEESARCEIKNTDFKALLIVKDAIRIDGGYILRASNVAGSKSFP 24637
QY 241 ----- 240
Db 24638 VNVKVLDRPPEGPVQVGTSEKSLTWSPPLODGGSDISHYVVEKRETSRLAWTVVA 24697
QY 241 ----- 240
Db 24698 SEVVTNSLKVTKLLEGNEYFRIMAVNKYGVGEPELESAPVLMKNPFLPGPPKSLVNTNI 24757
QY 241 ----- 240

Db	24758	AKDSMTVCWNRPSDGGSEIIIGYIIVEKRDGIRWIKCNKRITDLRLRVTLGTEDHEYE	24817
QY	241	-----	240
Db	24818	FRVSAENAGVGEPSPATVYKACDPVKPGPTNAHIVDTTKNSITLAWKPIYDGGSE	24877
QY	241	-----	240
Db	24878	ILGYVVEICKADBEEMQIVTPOTGLRVTRFEISKLEHQEYKIRVCALNKVGLGEATSP	24937
QY	241	-----	240
Db	24938	GTVPKPEDKLEAPELDLDSELRGIVVRAGGSARIHIPKGRPTPEITWSREGEFTDKVQ	24997
QY	241	-----	240
Db	24998	IEKGVNYTOLSDNCDNRNDAGKYLKLENSGSKSAFVTVKVLDTPGPPQNLAKEVRKD	25057
QY	241	-----	240
Db	25058	SAFLVWEPIIDGAKVKNYVIDKRESTRKAYANVSSCKTSFKVENLTEGAIYFRYM	25117
QY	241	-----	240
Db	25118	AENERGVGVVETDAVKAAPSPPGKVTLTDVQSATSASLMEKPEHDGSRVLGVYE	25177
QY	241	-----	240
Db	25178	MQPKGTEKWSIAESKVCNAVTVGLSSGOEYQFRVKAENEKSDPRVLGVVPIAKDLTI	25237
QY	241	-----	240
Db	25238	QPSLKLPFNTYSTQAGEDLKEIPVIGRPNISWVKDGEPLKQTRVNVEETATSTVLH	25297
QY	241	-----	240
Db	25298	IREGNKDDFGKYTVTATNSAGTATENLSVIVLEKPGPPVGVRFDEVSADFVVISWEPPA	25357
QY	241	-----	240
Db	25358	YTGCGQISNYIIVEKRDTTTTTHWMSAVVARTTIKIKLTKGTEYQFRIFAENRYGKSAP	25417
QY	241	-----	240
Db	25418	LDSKAVIVQYPFKEPGPGTPEVTSISKDQMLVQWHEPVDNGGTKIIGYHLEQKEKNSLT	25477
QY	241	-----	240
Db	25478	WVKLNKTPIDQTKFKTTGLDGLGEYEFKVSANIVIGIKPSKVSECFVARDPCDPPGRPE	25537
QY	241	-----	240
Db	25538	AIVITRNVTLWKKPAYDGGSKIYGYIIVEKKDLPDGRMKASFTNVLETFVTSGLVED	25597
QY	241	-----	240
Db	25598	QRYEFVIRIARNAAGSEPSDSSGALTARDEIDAPNASLDPKYKDVIVVHAGETFFLEAD	25657
QY	241	-----	240
Db	25658	IRGKPIPDVYWSKDGKELEETAARMEIKSTIQKTTLVVKDCIRTDGGQYILKLSNVGGTK	25717
QY	241	-----	240
Db	25718	SIPITVKVLDPRGPPGEPGLKVTGVTAEKCYLAWNPPLODGGANISHYIIEKRETSRLSWT	25777
QY	241	-----	240
Db	25778	QVSTEVALNYKVTKLLPCNEVIFRVMVKNYKIGEPLESGPVTACNPYKPPGPSTPEV	25837
QY	241	-----	240
Db	25838	SAITKDSMVVTVWARVPVDDGGTEIEGYILEKRDKEGVRWTKCNKKTLDLRLRVTLGTEGH	25897
QY	241	-----	240
Db	25898	SYEFRAAENAGVGEPSPSVYRACDALYPPGPPSNPKVTDTSRSSVSLAWSKPIYDG	25957
QY	241	-----	244
Db	25958	GAPVKGYYVEVKEAAADEWTTCTPTGLQGKQFTVTKLKENTEYNFRICAINSEGVEPA	26017
QY	245	-----	244
Db	26018	TLPGSVVAQERIEPPEIELDADLRKVVLRSATLRLFVTIKGRPEPEVKEKAEGILTD	26077
QY	245	-----	244
Db	26078	RAQIEVTSFTMLVIDNVTRPDSGRYNLLENNSGSKTAFVNVRLDSDSPAPVNLTIREV	26137
QY	245	-----	244
Db	26138	KKDSVTLISWEPLIDGGAKITNYIIVEKRETRKAYATITNNCTKTTFRIENLQEGCSYVF	26197
QY	245	-----	244
Db	26198	RVLASNEXGIGLPAETTEPVPKVSPPPLPPGRVTLVDVTRNTATIKWKPESDGGSKITGY	26257
QY	245	-----	244
Db	26258	VVEMOTKGSEKSTCTQVKTLEATISGLTAGEEYVFRVAANVEKGRSDPRQLGVPVIARD	26317
QY	245	-----	244
Db	26318	IEIKPSVELPFHTFNVKAREQLKIDVPFKRPOATVNMRRKDQTLKETTRVNVSSKVT	26377
QY	245	-----	244
Db	26378	SLSIKEASKEDVGYELCVSAGSITVPTIIVLDRPGPPGPIRIDEVSCSITISWNP	26437
QY	245	-----	244
Db	26438	PEYDGGCOISNVIIVEKKETTSTTHIVSOAVARTSIKIVRLTTGSEYQFRVCAENRYGKS	26497
QY	245	-----	244
Db	26498	SYESSAVVAEYPPFPSPGPGTKVVHATKSTMLVTWQVPVNDGGSRVIGYHLEYKERS	26557
QY	245	-----	244
Db	26558	ILWSKANKILITADTQMKVSGLDGLMYEYRYVAENIAGIGKCSKCEPVPARDPCDPPGQ	26617
QY	245	-----	244
Db	26618	PEVTNITRKSYSKWSKPHYDGGAKITGYIVERRELPGDRWLKCNNTNIOETTYFEVTELT	26677
QY	245	-----	244
Db	26678	EDQRYEFVRVARNADSVSESESTGPIIVKDDVEPPRVMDVKFROVIVVKAGEVLKIN	26737
QY	245	-----	244
Db	26738	ADIAGRPLPVISWAKDGIIEERARTEIISTDNHTLLTVKOCIRRDGTQGYVLT LKNVAGT	26797
QY	245	-----	244
Db	26798	RSVAVNCKVLDKPGPPAGPLEINGLTAEKCSLSWGRPQEDGGADIDYIIVEKRETSHLAW	26857
QY	245	-----	244
Db	26858	TICEGELQWTSCKVTLLKNGNEYIFRVTGVNKYGVGEPLSVAIKALDPFTVPSPPTSLE	26917
QY	245	-----	244
Db	26918	ITSVTKESWTLCSRPESDGGSEISGYIITERREKSLRWVRVNNKPKVYDLRVKSTGLREG	26977

QY	245	-----	244
Db	26978	CEVEYRYAENAGLSLPSLRAEDPVFLPSPSPKPIVDSGKTTITIAWVKPLFD	27037
QY	245	-----	244
Db	27038	GGAPITGYTVEYKSSDDTKTSIQSLRGTEYTGISLTTGAEYVFRVKSYNKVGASDPSPD	27097
QY	245	-----	244
Db	27098	SSDQIAKEREBEPLFDIDSEMRKTLIVKAGASFTMTVPRGRPVNVLSKPDLDLRT	27157
QY	245	-----	244
Db	27158	AYVDTTDSRTSLTIENANRNSGKYTLTIQNVLSAASLTILVVKVLDTPGPPTNITVQDVT	27217
QY	245	-----	244
Db	27218	KESAVLSWDVPENDGGAPVKNYHIEKREASKAWSVTNNCRLSYKVTNLQEGAIYFR	27277
QY	245	-----	244
Db	27278	VSGENERGVCIPAETKEGVKITEKPSPEKLGVTISKDSVSLTWLKPEDHGGSRIVHYV	27337
QY	245	-----	244
Db	27338	VEALEKGQKNWCAVAKSTHVVSGLRENSEYFFRFAENQAGLSDPRELLLPVLIKEQ	27397
QY	245	-----	244
Db	27398	LEPPEIDMKNFPHSTVYVRAGSNLKVDIPISGRPLPKVTLSDRGVPLKATMRFNTEITAE	27457
QY	245	-----	244
Db	27458	NLTINKESVTADAGRYEITAANSSGTTKAFINIVLDRPGPTGPVVISDITEESVTLK	27517
QY	245	-----	244
Db	27518	WEPPKYDGSQVTNYILLKRETSTAVWTEVSATVARTMMKVMKLTGEEYQFRIKAENRF	27577
QY	245	-----	244
Db	27578	GISHIDSACVTVKLPYTTGPPSPTPWTVNVTRESITVGNHPEVNSGSAVGVHLEMKD	27637
QY	245	-----	244
Db	27638	RNSILWOKANKLVIRTTHEKVTITISAGLIYEFVYAENAGVGKPSHPSEPLAIDACEP	27697
QY	245	-----AFD- 	247
Db	27698	PRNVRTIDISKNSVLSWOOPAFDGGSKIRGYIVERRDLPDGRWTKASFNTVTETQFIIS	27757
QY	248	-----	247
Db	27758	GLTONSOYEFVRFARNAGVSNPSEVVGPTICIDSYGGPVIDLPLEYTEVWKYRACTSV	27817
QY	248	-----	247
Db	27818	KLRAGISGKPAPTIIEWYKDDKELQTNALVCVENTDLDASILIKDADRLNSGCYELKLRNA	27877
QY	248	-----	247
Db	27878	MGSASATIRVQILDKPGPGGPIEFKVTVAEKITLLWRPPADGGAKITHYIVEKRETSR	27937
QY	248	-----	247
Db	27938	VWMSWSEHLEECIITTTKIKNEYIFRVRVAVNKYIGEPLESDSVVAKNAFVTPGPPG	27997
QY	248	-----	247
Db	27998	IPEVTKITKNSMTVWMSRPIADGSDISGYFLEKRDKKSILGWPKVLKETIRDTRQKVTGL	28057
QY	248	-----	247

Db	28058	TENSQYQYRVCANAAQGPFPSEPFYKAADPIDPPGPAKIRIADSTKSSITLGWSKP	28117
QY	248	-----	247
Db	28118	VYDGSASVTGYVVEIROGEEETVSTKGEVTRTEYVVVNLKPGVNYFRVSAVNCAGQ	28177
QY	248	-----	247
Db	28178	GEPIEMNEPVOAKDILEAPEILDVALRTSVIAKAGEDVQVLIPFKGRPPPTVTRKDEK	28237
QY	248	-----	247
Db	28238	NLGSARYSTENTDSSLLTIPOVTRNDTKYILTIENGVEPKSSTSVKVLDTPAACQ	28297
QY	248	-----	247
Db	28298	KLOVKHVSRTVTVLLWDPPLIDGGSPIINYVIEKRDATKRTWSVSHKCSSTSKLIDL	28357
QY	248	-----	247
Db	28358	EKTPFFRVLAEINEIGEGEPCETEPVKAAEVPAPIRDLMSKDKSTKTSVILSWTKPDPDG	28417
QY	248	-----	247
Db	28418	GSVITEYVWERKKGQETWSHAGISKTCIEVSQLEKQSVLEFRVFAKNEKGLSDPVTIG	28477
QY	248	-----	247
Db	28478	PITVKELIITPEVLDSDIPGAQVTVRIGHNVHLELYKGRPKPSISWLKDLPLUKSEFV	28537
QY	248	-----	247
Db	28538	RFSKTENKIITLSIKNAKKEHGGKYTVILDNAVCHIAVPIVITVLGPPSKPGPIRDEIK	28597
QY	248	-----	247
Db	28598	ADSVILSWDPEDNGGEITCYISIEKRETSQTNWKMVCSSVARTTFKVPNLVKAQYQFR	28657
QY	248	-----	247
Db	28658	VRAENRYGVSOPLVSSIIIVAKHQPRIPGPPGKPVYINVTSDGMSLTWDAPVYDGGSEVTG	28717
QY	248	-----	247
Db	28718	FHVEKKERNILMQKVNTSPISGREYRATGLVEGLDYQFRVYAENSAGLSSPSDPSKFTL	28777
QY	248	-----	247
Db	28778	AVSPVDPPTGPDYIDVTRETITLKNPPLRDGSGKIYGVYSIEKROGNERWRCNFTDVSE	28837
QY	248	-----	247
Db	28838	COYTVTGLSPGDRYEFRIIARNNAVGTISPPSSQSGIIMTRDENVPPIVEFGPEYFDGLII	28897
QY	248	-----	247
Db	28898	KSGESLRKALVQGRPVPRVTFWKDGVIEKRMNMEITDVLGSTSLFVRDTRDRHGVYT	28957
QY	248	-----	247
Db	28958	VEAKNASGAKAEIKVKVQDTPGKVVGPPIRFTNITGEKMTLWMDAPLNDGCAPITHYIIE	29017
QY	248	-----	247
Db	29018	KRETSRLAWALIEDKCEAQSYTAKLINGNEYQFRVSAVNVKFGVGRPLDSDPVVAQIOYT	29077
QY	248	-----	247
Db	29078	VPDAPGIPESNITGNSITLTWARPESDGGSEIOQYILERREKKSTRWVKVSKRPISET	29137
QY	248	-----	247

Db 29138 RFKVTGLTEGNEYEFHVMAENAGVPASGISRLIKCREPVNPPGPTVVKVTDTSKTTV 29197
QY 248 ----- 247
Db 29198 SLEWSKPVFDGMEIIGYIEMCKADLGDWHKVNAAECVKTRYTVDLQAGEYKFRVSA 29257
QY 248 ----- 247
Db 29258 INGAGKDSCEVTGTIKAVDRUTAPELDIDANFKOTHVVRAGASIRLFAYQGRPTTAV 29317
QY 248 ----- 247
Db 29318 WSKPDSNLSLRADIHTDSFSTLTVENCNRNNDAGKYTLTVENNSGKSTITFTVKVLDTPG 29377
QY 248 ----- 247
Db 29378 PPGPITFDKVRGTSATLMDAPLLDGGARIHHYVVEKREASRSQWVISEKCTRQIFKVN 29437
QY 248 ----- 247
Db 29438 DIAEGVPYFRVSAVNEYGVEPEPEIVATEQPAPPRLLDVDTSKSSAVLAWLKPD 29497
QY 248 ----- 247
Db 29498 HDGSRITGYLLEMRQKGSDFWEAGHTKQLTFTVERLVEKTEYEFVRKAKNDAGYSEPR 29557
QY 248 ----- 247
Db 29558 EAFSSVLIKEPOIEPTADLTGITNOLITCKAGSPFTIDVPTSGRPAPKVTHKLEEMRLKE 29617
QY 248 ----- 247
Db 29618 TDRVSITTTKDRTTLTVDKSMRGDSGRYFLTLENTAGVTFSTVVWIGRPGVTGPTEV 29677
QY 248 ----- 247
Db 29678 SSVSAESCVLSWGEKPGDGGTEITNYIVEKRESGTTANOLVNSSVKRTQIKVTHLTKYME 29737
QY 248 ----- 247
Db 29738 YSFRYSSENREGVKPLESAPIIAEHFPVPSAPTRPEVYHVSANAMSIRWEEPHDGS 29797
QY 248 ----- 247
Db 29798 KIIGYWEKKEKRTILWVKENKVPCLCENYKVTGLVEGLEYQFRTYALNAAAGVSKASEAS 29857
QY 248 ----- 247
Db 29858 RPIMAQNPVDAGRPEVTDVTRSTVSLIWSAPAYDGGSKVGVYIIRKPKVSEVGDGRWLK 29917
QY 248 ----- 247
Db 29918 CNYTIVSDNFTVTALSEGDYEFVRVLAKNAGVISKSESTGPVTCRDEYAPPKAELDA 29977
QY 248 ----- 247
Db 29978 RLHGLDVTIRAGSLDLDAVGGKPEPKIITWTKGDKELDCEKVSLOYTKRATAVAKFC 30037
QY 248 ----- 247
Db 30038 DRSDSGKTYLTVKNASGTVKAVSMVKVLDSPGCGKLTVSRTQEKTLAWSLPQEDGGA 30097
QY 248 ----- 247
Db 30098 EITHYIVERRETSLRNWVIVEGECTLSYVVYTRLIKNEYIFRVRANKYGPVGVSEPE 30157
QY 248 ----- 247
Db 30158 IVARNSFTIPSPGPIEVEGTGKEHIIQWTKPESDGGNEISNYLVDKREKSLRWTRVN 30217
QY 248 ----- 247
Db 30218 KDYVVYDTRLKVTSLMGCDYQFRVTAVNNAAGNSEPSEASNFISCREPSYTPGPPSAPRV 30277

QY 248 -----WLKAFYD----- 254
Db 30278 VDTTKHSISLAWTKPMYDGGDIVGYVLEMOEKDQDQWVRVHTNATIRNTEFTVPDLKMG 30337
QY 255 ----- 254
Db 30338 OKYSFRAAVNVKMGSEYSESAIEIEPVERIEIPDLELADLLKKTVTIRAGASRLMVSV 30397
QY 255 ----- 254
Db 30398 SGRPPPVITWSKQIDILASRAIIDTTESYSLLIYDKVNRDAGKYTIEAENOSGKKSATV 30457
QY 255 ----- 254
Db 30458 LVKYVDTPGCPSPVKVKEVSRDSVTITWEIPTIDGGAPVNNYIVEKREAAAMRAFKVTITK 30517
QY 255 ----- 254
Db 30518 CSKTLIRISGLVEGTMYYFRVLVLPENIYGIGPECTSDAVLSEVPLPAKLEVVDTKST 30577
QY 255 ----- 254
Db 30578 VTLAWEKPLYDGGSKLTGYVLEACKAGTERMKVVTLPKTVLEHTVTSLSNEGEQYLFRI 30637
QY 255 ----- 254
Db 30638 AQNEKGVSEPRETVATVQDLRVLPDIDLTSTMPQKTIHPAGRPVELVPIPIAGRPPAA 30697
QY 255 ----- 254
Db 30698 SWFTAGSKLRESERVVETHTKVAKLTIRETTIRDTGEYTLLELKNVTGTTSETIKVIIL 30757
QY 255 ----- 254
Db 30758 KPGPTGPIKIDEIDATISITISWEPPELDGGAPLSGYVVEQDRAHPGWLVPVSEVTRST 30817
QY 255 ----- 254
Db 30818 FKFTRLTEGNEYVFRVAATNRFIGSYLQSEVIECRSSIRIPGPPETLIQIFDVSRDGNTL 30877
QY 255 ----- 254
Db 30878 TWYPPEDDGGSQVTGYIVERKEVRADRWVRVNVKVPVTMYRSTGLTEGLEVEHVRVTAIN 30937
QY 255 ----- 254
Db 30938 ARGSGKPSRPSKPIVAMDPIAPPKQPONPRVTDTRTSVSLAWSVPEDEGSKVTGYLIE 30997
QY 255 -----KVAEK 259
Db 30998 MQKVDQHEWTKCNTPTPKIREYTLTHLPQGAERYFRVLACNAGGPEAEVPGTVKVTM 31057
QY 260 L----- 260
Db 31058 LEYPDYELDERYQEGIFVRQGVIRLTIPIKGPFPICKWTKEGQDISKRAMIATSETH 31117
QY 261 ----- 260
Db 31118 ELVITKEADRGDSGYDLVLENKCKKAVYIKVRVIGSPNSPEGPLEYDDIQVRSVRVSWR 31177
QY 261 ----- 260
Db 31178 PPADGGADILGYILERREVPKAAWYITDSRVROTSLVVKGLENVEHFRVSAENQFI 31237
QY 261 ----- 260
Db 31238 SKPLKSEBPVTPKTPLANPPEPPSNPPEVLDVTKSSVLSWSRPKDDGGSRTVGYIERKE 31297
QY 261 ----- 260
Db 31298 TSTDKNVRHNKTOITTTMYVTGLVPAEYQFRIIAQNDVGLSETSPASEPVCKDPDK 31357


```
QY 261 ----- 260
Db 31358 PSQGELEILSISKDSVTLOWEXPECDGGKEILGYWEYRQSGDSAWKSKNERIKDKQF 31417
QY 261 ----- 260
Db 31418 TIGGLEATEYEFVFAENETGLSRPRTAMSIKTLTSGEAPCIRKEMKDVTTKLGEAA 31477
QY 261 ----- 260
Db 31478 QLSQIVGRPLDIKWYRFKELIQSRKYMSSDGRTHLTVMTEQEDGVVTCIATNE 31537
QY 261 ----- 260
Db 31538 VGEVETSKLLQATPOFPGYPLKEKYGAVGSTLRHVMYICRPVPMATWFGKLLQ 31597
QY 261 ----- 260
Db 31598 NSENITIENTHYTHLYMKNVQRTHAGKYKVOLSNVFGTVDAILDVEIQDKPKPTGPI 31657
QY 261 ----- 260
Db 31658 VIEALLKNSAVISWKKPPADGGSWITNYVVEKCEAKEGAEWQLVSSAISVTTCRIVNLTE 31717
QY 261 ----- 260
Db 31718 NAGYFPRVSAQNTFGISDPLEVSSVVIKSPFEKPGAPKPTITAVTKDSCVVAVKPPAS 31777
QY 261 ----- 260
Db 31778 DGGAKIRNYYLEKREKONKWI SVTTEIRETVFSVKNLIEGLEYPFRVCENLGGESEW 31837
QY 261 ----- 260
Db 31838 SEISEPTPKSDVPIQAPHEKELNLRVYQSNATLVCKVTHGHPKPIVWYRQGEIIA 31897
QY 261 ----- 260
Db 31898 DGLYRIQFKGVHQLIIASVTDATVYQVRATNNGSGSVGTASLEVEVPKAIHLPKT 31957
QY 261 ----- 260
Db 31958 LEGMAVHALRGEVWSIKIPFGKPDVITWQKQDLNNGHYQVIVTRFTSLVFPNG 32017
QY 261 ----- 260
Db 32018 VERKAGFYVVCANRIGIDQKTVELDVADVPDPPRGVKYSDVSRDSVNLTWTEPASDGG 32077
QY 261 ----- 260
Db 32078 SKITNYIVEKCATTAERWLRVQARETRYVINLFGKTSYQFRVIAENKFGLSKPSPE 32137
QY 261 ----- 260
Db 32138 PTITKEDKTRAMYDEEDVETREVSMTKASHSTKELYEKYIMTAEDLGRGFGIVHRCVE 32197
QY 261 ----- 260
Db 32198 TSSKITYMAKFVKVKTQDQVLVKKEISILNIARHNILHLHESFESMEELVMIFEISGL 32257
QY 261 ----- 260
Db 32258 DIFERINTSAFELNREIVSVHVCEALQFLSHNIGHFDIRPENIYQTRRSSTIKII 32317
QY 261 ----- 260
Db 32318 EFGQARQLKPDNRLFTTAPYYAPEVHQHDVVSTATDMWSLGLTVYLLSGINPF LAE 32377
QY 261 ----- KEAF 264
:||||
Db 32378 TNOQIENIMNAEYTFDEEAF 32398
```

```
RESULT 14
Q10466 PRELIMINARY; PRT: 26926 AA.
ID Q10466;
AC Q10466;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Titin, heart isoform N2-B (EC 2.7.1.-) (Connectin).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HEART;
RX MEDLINE=96026330; PubMed=7569978;
RA Labeit S., Kolner B.;
RT "Titins: giant proteins in charge of muscle ultrastructure and
RT elasticity.";
RL Science 270:293-296(1995).
RN [2]
RP SEQUENCE OF 2277-25376 FROM N.A.
RX MEDLINE=92258380; PubMed=1582406;
RA Labeit S., Gautel M., Lakey A., Trinick J.;
RT "Towards a molecular understanding of titin.";
RL EMBO J. 11:1711-1716(1992).
RN [3]
RP SEQUENCE OF 1976-2014 FROM N.A.
RA Labeit S.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=95331314; PubMed=7607248;
RA Gautel M., Castiglione-Morelli M.A., Pfuhl M., Motta A., Pastore A.;
RT "A calmodulin-binding sequence in the C-terminus of human cardiac
RT titin kinase.";
RL Eur. J. Biochem. 230:752-759(1995).
CC -1- FUNCTION: THIS GIANT MUSCLE PROTEIN MAY BE INVOLVED IN MUSCLE
CC ASSEMBLY AND IN MAINTAINING THE STRUCTURAL INTEGRITY OF
CC SARCOMERES. MAY HAVE PROTEIN KINASE ACTIVITY.
CC -1- ALTERNATIVE PRODUCTS: A NUMBER OF FORMS OF THIS PROTEIN ARE
CC PRODUCED BY ALTERNATIVE SPLICING WHICH DIFFER IN TISSUE
CC DISTRIBUTION. DIFFERENT SIZE TRANSCRIPTS MAY ALSO EXIST WITHIN ANY
CC ONE TISSUE. THE SEQUENCE SHOWN HERE IS THAT OF THE HEART ISOFORM
CC N2-B.
CC -1- TISSUE SPECIFICITY: MUSCLE-SPECIFIC.
CC -1- SIMILARITY: TO THE CATALYTIC DOMAINS OF OTHER SERINE/THREONINE
CC KINASES.
CC -1- SIMILARITY: BELONGS TO IMMUNOGLOBULIN SUPERFAMILY. CONTAINS 112
CC IMMUNOGLOBULIN C2-LIKE DOMAINS AND 132 FIBRONECTIN TYPE III-LIKE
CC DOMAINS.
DR EMBL; X64698; CAA45939.1; -
DR EMBL; X83270; CAA58243.1; -
DR EMBL; X64697; CAA45938.1; -
DR EMBL; X90568; CAA62188.1; -
DR EMBL; X64699; CAA45940.1; -
DR HSSP; P56276; ITLK
DR InterPro; IPR000282; Cytok_receptor_2.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR000577; FGGY_kin.
DR InterPro; IPR003962; FNIII_repeat.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR000129; Peptidase_S24.
DR InterPro; IPR02016; Peroxidase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00041; fn3; 132.
DR Pfam; PF00047; Ig; 91.
DR Pfam; PF00069; pkinase; 1.
```


Qy 39 -----EUTC----- 42
Db 1718 YGVAYSRDSGIITCRATNKYGTDHTSATLIVKDEKSLVEESOLPEGRKGLRIBELERMA 1777
Qy 43 ----- 42
Db 1778 HEGALTGVTTDQKEKQKPDIVLYPEPVRVLEGETARFCRVGTGYPQPKVNWYLANGOLIRK 1837
Qy 43 ----- 42
Db 1838 SKRFRVRVDGIHYLDIVDCKSYDTGCVKVTAEENPEGVIEHKVKLEIQOREDFRSVLRRAP 1897
Qy 43 ----- 42
Db 1898 EPRPEFHVHEPKLQFEVQKVDPRVDTTETKEVVVKLRAERITHKEKVPSESEELRSKFR 1957
Qy 43 ----- 42
Db 1958 RTEGYYEAITAVELSKRKKDESYELLRTKDKELLHWTKEELKALAEKGKITIPT 2017
Qy 43 ----- 42
Db 2018 FKPKIELSPMEAPKIFERIQSQTQVGSDAHERVRVVGKPDPECEWYKNGVKIERSDR 2077
Qy 43 -----VVVDVSHED----- 51
Db 2078 IYTWPNEDNVCLEIWDVTAEDSASIMVKAINIAGETSSHAFLLVQAQOLITFTQELQDV 2137
Qy 52 -----PEVKFNHYVDGVEVH----- 66
Db 2138 VAKEKDTMATFECETSEFPFVKVYKDGMEVHEGDKYRMHSDRKHVHFLSILITDTSDAED 2197
Qy 67 -----NAKT----- 70
Db 2198 YSCVLVEDENKTTAKLIVEGAVVEFKELQDIEVPESYSCELECIVSPENIEGKWYHND 2257
Qy 71 -----KPR----- 73
Db 2258 VELKSNKGYTTITRRGRQNLTKVDTKEDQGEYSFVIDGKTTCKLKMKPRPIALQGLS 2317
Qy 74 ----- 73
Db 2318 DQKCEGDIQVLEKVKVLSLESVEGWMKDGQEVQPSDRVHIIVDKOSHMLLIJEDMTKEDAG 2377
Qy 74 ----- 73
Db 2378 NYSETIPALGLSTGRVSVYSDVITPLKDVNVIEGTAKVLECKVSPDVTSVKWLNDE 2437
Qy 74 ----- 73
Db 2438 QIKPDDRQVAIVKGTQKRLVINRTHASDEGYPYKLIIVGRVETNCNLSVEKIKIIRGLRDLT 2497
Qy 74 ----- 73
Db 2498 CTETQNVVFEVLSHGIDVLWNFKDEIKPSKRYKIEAHGKIYKLTVLNMKDDGKRYT 2557
Qy 74 ----- 73
Db 2558 FYAGENMTSGKLTAVAGGAISKPLTDQTVAESQEAFFECEVANPDSKGEWLKDLPLTN 2617
Qy 74 ----- 73
Db 2618 NIRSSEDGHRRLIIAATFKLDDIGEYTYKVATSKTSAKLKVEAVKIKKTLKLNLTJVTETQD 2677
Qy 74 ----- 73
Db 2678 AVFTVELTHPNVKGWQVIKNGWLVLESNEKYAISVKGTIYSLRIKNAIVDESIVGFRIGR 2737
Qy 74 ----- 73
Db 2738 LGASARLHVETVKTIKKPKDVTALENATVAFEVSVSHDTPVPKWFHKSVEIKPSDKHRLV 2797
Qy 74 ----- 73

Db 2798 SERKVHKLMLQNISPSDAGEYTAVVQGLECKAKFLVETLHITKTMKNIEVPETKTASPEC 2857
Qy 74 ----- 73
Db 2858 EVSHFNVPMSWLKNGVEIEMSEKFKIVVQGLHQLIIMNTSTEDSAEYTFVCGNDQVSAT 2917
Qy 74 ----- 73
Db 2918 LTVTPIMTISMLKDINAEEKDTITFEVTNVYEGISYKWLKNGVEIKSTDQCMRTKKLTH 2977
Qy 74 ----- 73
Db 2978 SLNIRNVHFGDAADYTFVAGKATSTATLYVEARHIEFRKHIDIKVLEKKRAMFECEVSE 3037
Qy 74 ----- 73
Db 3038 PDITVQMMKDOELOITDRIKIQEKYVHRLIIPSTRMSDAGKYTVVAGGNVSTAKLFE 3097
Qy 74 ----- 73
Db 3098 GRDVRIRSIKKEVQVIEKQRAVVEFEVNEDDVDAHMYKDGIEINFQVOERHKYVVERRIH 3157
Qy 74 ----- 73
Db 3158 RMFISETRQSDAGEYTFVAGNRSSVTLYVNAPEPPVQLQELQVTVQSGKPARFCAMIS 3217
Qy 74 -----EE----- 75
Db 3218 GRPQPKISWYKEBQLLSTGFKCKFLHDQGEYTLILLIEAFPEDAAVYTCEAKNDYGVATTS 3277
Qy 76 ----- 75
Db 3278 ASLSVEPEVVPDQEMPVYPPIITPLQDVTSEGPQRCRVSGTDLKVSWSKDKK 3337
Qy 76 -----QYNSTYRV----- 83
Db 3338 IKPSRFRMTQFEDTYQLEIAEAYPEDEGTVTFVANNVAVQSVSTANLSLEAPESILHER 3397
Qy 84 ----- 83
Db 3398 IEQIEIEMEMKFFSSFLSABEEGLHSAELQSLKINETLELLESVPYPTKFDSEKEGTGP 3457
Qy 84 -----VSVL----- 87
Db 3458 IFIKEVSNADISMGDVATLSVTVIGIPKPIQWPFNGVLLTPSADYKFVFDGDDHSLIIL 3517
Qy 88 ----- 87
Db 3518 FTKLEDEGEYTCMASNDYKGTICSAYLTKINSKGEGHKDTETESAVAKSLEKLGPCPPHF 3577
Qy 88 -----TVLH----- 91
Db 3578 LKELKPIRCAOGLPAIFEYTVVGEPAPTVTWFKENKOLCTSVYTIHHPNGSGTFFIVND 3637
Qy 92 ----- 91
Db 3638 POREDSGLYICKAENMLGESTCAAEILLVLEDDTMDTDPCKAKSTPEAPDPQTPKGP 3697
Qy 92 ----- 91
Db 3698 AVEALDSEIATFVKDTILKAALITEENOQLSYEBIAKANELSSOLPLGAQLQSILEQ 3757
Qy 92 ----- 91
Db 3758 DKLTPESTREFLCINGSIHQFPLKPEPSNLOLQIVOSOKTFSKEGILMPEEPETOAVLSD 3817
Qy 92 ----- 91
Db 3818 TEKIPPSAMSTEQINSLTVEPLKTLAEPENYPOSSIEPPMHSYLTSAEVEVLSLKCKT 3877
Qy 92 ----- 91

```
Db 3878 VSDTNRQVRTLOKQEAQSALILSQAEGHVESLQSPVMISQVNYEPLVPSEHSCTEG 3937
QY 92 ----- 91
Db 3938 GKILIESANPLENAGDSAVRIEKGSLRPLALEKQVLLKEEHSNDNVMPDQIIESK 3997
QY 92 -----ODWL- 95
Db 3998 REPVAIKKVOEGRDLSKESLLSGIPEQRNLNKIQICRALQAASQPGLFSEWLR 4057
QY 96 ----- 95
Db 4058 NIEKVEAVNITQEPHIMCYLVTSKASVTEVTIIBEDVPOMANLKMELRDALCAI 4117
QY 96 ----- 95
Db 4118 IYBEIDILTAEGPRIOGAKTSLQEMDSFGSQKVEPITEPEVESKYLISTEVSFYFNV 4177
QY 96 ----- 95
Db 4178 QSRVKYLDATPVTKGVASAVVDEKQDESLKPSEKESSSESGTEEVATVKIQEABGGL 4237
QY 96 ----- 95
Db 4238 IKEDGPMIHTPLVDTVSEGDIVHLTTSITNAKEVNMVYFENKLVPSDEKFKCLODQNTYT 4297
QY 96 -----NGK- 98
Db 4298 LVIDKVNTEDHQGEYVCEALNDSGKTATSAKLTJVVRAAPVIKRKTEPLEVALGHLAKFT 4357
QY 99 -----EYKCKVSN- 106
Db 4358 CEIQSAPNVRFWFKAGRIEYSDKCSIRSSKYISSLEILRTQVDCGEYTCASNEYGS 4417
QY 107 ----- 106
Db 4418 VSCATLTVTPGGEKKVRKLLPERPEKPEEVVLKSVLRKRPEEPEKPKLEKVKX 4477
QY 107 ----- 106
Db 4478 PAVPEPPPKPVEVEVPTVTKRERKIPEPTKVPEIKPAIPLPAPPEKPKPAEVKTIP 4537
QY 107 -----KALPAPIE- 114
Db 4538 PPVEPETPTAAPVTVVVGKKAABAKPAKEAAKPGIKGVPKTPSPTEARRKLPG 4597
QY 115 ----- 114
Db 4598 SGGEKPPDEAPFTYQLKAVPLKFVKKEIKDIIILTESEFVGSSEAFCELVSPSTAITWMKD 4657
QY 115 ----- 114
Db 4658 GSNIRESPKHRFTADGDKRKLHIIDVLSAGEYTCVLRGNKEKTSTAKLVVEELPVRF 4717
QY 115 -Ktis- 118
Db 4718 VKTLEEEVTVVGQPLYLSCELNKRDEVVVRKDKGIWEKPGRIVPGVIGLMRALTI 4777
QY 119 ----- 118
Db 4778 DDTDAGTYTVVENANNLESCSCVKVVEVIRDLVLPKPIRDQHVKPGKTAIFACDTAKDTP 4837
QY 119 ----- 118
Db 4838 NIKWFGYDEIPAEPNDKTEILLRDNHLYLKIKNAMPEDIAEVAVEIEGKRYPAKLTIGE 4897
QY 119 ----- 118
Db 4898 REVELLKPIEDVTIYEKESASFAEISEADIFQWMLKGELLRPSPTCEIKAEAGGKRELT 4957
QY 119 ----- 118
Db 4958 LHKVKLDQAGEVLYQALNATITAILTVKEILDFAVPLKDVTPPERRQARFECVLTREAN 5017
```

```
QY 119 ----- 118
Db 5018 VIWSKGPDIKSSDKFDIADGKKHILVINDSOFDDDEGVYTAEEGKTSARLFTVGTGIRL 5077
QY 119 ----- 118
Db 5078 KPMSPLEDDQTVKEGETATFVCELSHEKMHVWFKNDAKLHSTRTVLISSEBKGTHKLEMKE 5137
QY 119 ----- 118
Db 5138 VTLLDDISQIAQVKELSSSTAQLKVLADPYFTVKLHDKTAVEKDEITLKEVSKDVPVKW 5197
QY 119 ----- 118
Db 5198 FKDGEIIVSPKYSIKADGLRLIKKADLKDKEVVCDCGTDKTKANVTVEARLIEVE 5257
QY 119 -----KAGQOP- 124
Db 5258 KPLYGVEVFGETAHFEIELSEDPVHGQWKLKGOPLTASPCDEIIEDGKKHILHNCOL 5317
QY 125 -----REPO- 128
Db 5318 GMTGEVSFOAANAKSAANLKVKEPLIFITPLSDVKVFEKDEAKFECEVSRPKTPRWLK 5377
QY 129 ----- 128
Db 5378 GTQEIITGDDRFELIKDGTKHSMVKSAAFEDEAKYMFEAEDKHTSGKLIIEGIRLKLFTP 5437
QY 129 ----- 128
Db 5438 LKDVTAKEKESAVFTVELSHDNIRVKWFKNDQRLHTRTSVSMODEGKTHSITFKDLSIDD 5497
QY 129 ----- 128
Db 5498 TSOIRVEAMGMSSEAKLTVLEGDPYFTGKLQDYTGVEKDEVILQCEISKADAPVKWFKDG 5557
QY 129 ----- 128
Db 5558 KEIKPSNAVIKTDGKKRMLLKALKSDIGQYTCDCGTDKTSGLDIEDREIKLVRPLH 5617
QY 129 ----- 128
Db 5618 SVEVMEETARPETEISEDDIHANMKLGEALLQTPDCEIKEEGKTHSLVHLNCRLDQGTG 5677
QY 129 ----- 128
Db 5678 GYDFQAANYKSSAHLRVKPRVIGLLRPLKDVTVTAGETATFDCELSYEDIPVENYLLKGKK 5737
QY 129 ----- 128
Db 5738 LEPSDKVVPVPRSEGVHTLTLRDVKLEDAGEVOLTAQDFKTHANLFWKEPPEVFTKPLEDQ 5797
QY 129 ----- 128
Db 5798 TVEEGATAVLECEVSRENKAVKFKNGTEILSKYKYEIVADGVRKRLVIHDCPTPEDIKTY 5857
QY 129 ----- 128
Db 5858 TCDAKDFKTSCLNVVPPHVEFLRPLTDLOVREKEMARFECELSRENKAVKFKDGAETK 5917
QY 129 ----- 128
Db 5918 KGKKYDIISKGAVRILVINKCLDDDEAEYSCVARTARTSGMLTVLEEEAVFTKNLANIEV 5977
QY 129 ----- 128
Db 5978 SETDTIKLVCEVSKPAEVIWYKGDDEIETGRIEILTEGRKRIILVIONAHLEDAGNYNC 6037
QY 129 ----- 128
Db 6038 RLPSSRTDGVKVKHELAAEFISKPONLEITGEKAEFVCSISKESEFPVQWKRDDKTLESG 6097
```

Qy	129	-----	128
Db	6098	DKYDIADGKRVLVWKDQTLQDMGYVMVGAARAAHLTVIEKLRIVVPLKDRVKEQ	6157
Qy	129	-----VYTL-----	132
Db	6158	QEVVFNCEVNTGAKAKWFRNEAIFDSSKYYIILOKDLVYTLRIRDAHLDDQANYNVSLT	6217
Qy	133	-----	132
Db	6218	NHRGENVKSANLIVEBEDLRIVEPLKDIETMEKKSVTFWCKVNRNLNVTLKWTKNGEVP	6277
Qy	133	-----	132
Db	6278	FDRVSVRVKDKYKHLTIKDCGFPDEGEYIVTAGQDKSVAELIIIEAPTEFVEHLEQTV	6337
Qy	133	-----	132
Db	6338	TEFDVAVFSCQLSREKANVWYRNGREIKEGKYKFEKDGSIHRLIIKDCRLDDECEYAC	6397
Qy	133	-----	132
Db	6398	GVEDKSRARLFVEEIPVEIIRPPQDILEAPGADVFLAELNKDKVEVQWLRNNWVVQG	6457
Qy	133	-----	132
Db	6458	DKHQMSEGKIHLRQICDIKPRDQGEYRFIAKDKEARAKLELAAAPKIKTADQDLVVDVG	6517
Qy	133	-----	132
Db	6518	KPLTMVVPYDAYPKAAEWFENEPLSTKTIIDTTAEQTSFRILEAKRGDKRYKIVLQNK	6577
Qy	133	-----	132
Db	6578	HGAEGFINLKVIDVPGVRNLEVTFETFDGEVSLAWEEPLTDGSGKIIGYVYVVERRDIKR	6637
Qy	133	-----	132
Db	6638	TWVLATDRAESCEFTVTGLQGGVEYLFVRSARNRVCTGEPVEDNPVEARSKYDVPGPP	6697
Qy	133	-----	132
Db	6698	LNVTITDVRNFGVSLTWEPEYDGAETINYVIELRDKTSIRWDTAMTVRAEDLSATVTD	6757
Qy	133	-----	132
Db	6758	VVEGQYSFVRVRAQNRIGVGKPSAATPFVKVADPIERPSPVNLTSDDQTQSSVOLKWEP	6817
Qy	133	-----	132
Db	6818	PLKDGSPILGYIIERCEEKGNWIRCNMKLVPELTYKVTGLEKGNKYLYRVAENKAGV	6877
Qy	133	-----	132
Db	6878	SDPSEILGLTADDAFVEPTMDLSAFKDGLEVIPNPITILVPSTGYPRPTATWCFGDKV	6937
Qy	133	-----	132
Db	6938	LETGDRVMMKTLAYAEVLISPSERSDKGIYTLKLENRVKTIISGEIDVNVIARPSAPKEL	6997
Qy	133	-----	132
Db	6998	KFGDITKDSVHLTWEPPDDDGSPLTGCVVVEKREVSRKWTWKVMDFTDLEFVTPDLVQG	7057
Qy	133	-----	132
Db	7058	KEYLFKVCARNKCGEPAYVDEPNMSTPATVPDPENVKWRDRTANSIFLTWDPKND	7117
Qy	133	-----	132
Db	7118	GGSRIGYIIVERCPRGSKWVACGEPVAETKMVEVTGLEGKWYAYRVKTLNRQGAKPSR	7177
Qy	133	-----	132

Db	7178	PTEIQAVDTQEADEIEFLDVKLLAGLTVKAGTKIELPATVIGKPEKITWTAKMILKQD	7237
Qy	133	-----	132
Db	7238	KRITIENVPKSTVTIIVDSKRSDTGTYYIEAVNVCGRATAVVEVNVLDKPGPAAFDITD	7297
Qy	133	-----	132
Db	7298	VTNESCILLTNPPRDDGSGKITNYVVERRATDSEWHKLSSTVKDTNFKATKLIPNKEYI	7357
Qy	133	-----PPSR---DELTKNOVSLT-----	147
Db	7358	FRVAENNYGAGEDVQASPIAKYQFDPGPGPTRLEPSDITKDAVTLTWCEPDDDDGGSPI	7417
Qy	148	-----	147
Db	7418	TGYWVERLDPDTDKWRCNKMPVKDTTYRVKGLTNKKYFRVLAENLAGPKPSKSTEP	7477
Qy	148	-----	147
Db	7478	ILIKDPIIDPPWPGKPTVKDVGKTSVRLNWTKPEHDGGAKIESYVIEMLKTGTDEWVRVA	7537
Qy	148	-----	147
Db	7538	EGVPTTQHLLPGLMEGQEYSFRVRAVNKAGESEPSPDPVLCREKLYPPSPPRWLEVIN	7597
Qy	148	-----	147
Db	7598	ITKNTADLKWTPEKDGSGPITNYIVEKRDVRRKGWQTVDTTVKDTKCTVTPLEGSLYV	7657
Qy	148	-----	147
Db	7658	FRVAENNAIGOSDTEIEDSVLAKDFTTTPGPPYALAVDVTKRHRVLDKWEPPKNDGGRP	7717
Qy	148	-----	147
Db	7718	IQRVVEKKERLGRWVKAGTAGDCNFRVTDVIEGTEVOFQVRAENEAGVHGPSEBTE	7777
Qy	148	-----	147
Db	7778	ILSIEDPTSPSPPLDLHVTDAGRKHIAIAWKPKPEKNGSPIIGYHVEMCPVGTEKWMRV	7837
Qy	148	-----	147
Db	7838	NSRPIKDLKFVEGVVDPDKEYVLRVRAVNAIGVSEPSISENVVAKDPDCKPTIDLETH	7897
Qy	148	-----	147
Db	7898	DIIVIEGEKLSIPVPFRAVPVPTVSWHKDGKEVKASDRLTMKNDHISAHLEVPKSVRADA	7957
Qy	148	-----	147
Db	7958	GIYTTITLENKLSATASINVKIVGLPGPCKDIKASDITKSSCKLTWEPPEPDGTPILHY	8017
Qy	148	-----	147
Db	8018	VLERREAGRRTYIPVMNGENKLSWTVDLIPNGEYFFRVKAVNKVGGEYIELKNPVAQ	8077
Qy	148	-----	147
Db	8078	DPKQPPDPVDEVHNPTAEAMTITWKPLYDGGKIMGYIIIEKIAKEERWKRCKNEHLV	8137
Qy	148	-----	147
Db	8138	PILTYTAKLEEGKEYQFRVRAENAAIGISEPSRATPPTKAVDPIDAPKVILRTSLEVRKG	8197
Qy	148	-----	147
Db	8198	DEIADASISGSPYPTITWIKDENNVIVPEEIKKRAAPLVRRRRKGVEEPEFVLPLORL	8257
Qy	148	-----	147

Db 8258 SIDNSKGSQSLVRDSLRDPDHLGYMIKVENDHGIKAPCTVSVLDTGPPINFVFEDIR 8317
QY 148 ----- 147
Db 8318 KTSVLCKWBPPLDDGGSEIINYITLXKDKTKPDSEHIVVTSTLRHCKYSVTKLIEGKEYL 8377
QY 148 ----- 147
Db 8378 FRVRAENRFGPGPCVSKPLVAKDPGPDADPKPIVEDVTSNSMLVKWNEPKDNGSPIL 8437
QY 148 ----- 147
Db 8438 GYWLKREVNSTHWSRVNKSLLNALKANDVGLLEGLTYFVRVCAENAGRGFSPPDPK 8497
QY 148 ----- 147
Db 8498 TAHDPISPGPPIPRVTDTSSTTIELEWEPFAFNGGGEIVGVFDVKQLVGTNKNWSRCTEK 8557
QY 148 ----- 147
Db 8558 MIKVRQYTVKEIREGADYKLRVSANVNAAGEPPGETQPVTVAEPPQEPAPAVELDVSKGI 8617
QY 148 ----- 147
Db 8618 QIMAGKTLRIPAVVTGRPVPTKVWTKKEGELDKDRVIDNVGTSKSELIITKDALKRDHGRY 8677
QY 148 ----- 147
Db 8678 VITATNSCGSKFAARVEFDVPGPVLDLKPVTNRKMKLLNWSDEDDGGSEITGFIIE 8737
QY 148 ----- 147
Db 8738 RKDAKMTWRQPIETERSKCDITGLLEGQYKFRVIAKNKFCGPPVEIGPILAVDPLGP 8797
QY 148 ----- 147
Db 8798 PTSPEBLTYTERQSTITLDWKEPRNSGGSPIQGYIIERRHDKPDFERVNKRCLCPTTSF 8857
QY 148 ----- 147
Db 8858 LVENLDEHOMYEFVRKAVNEIGESEPSLPLNVVQDDEVPPPIKRLSVRGDTIKVAGE 8917
QY 148 ----- 147
Db 8918 PVHIPADVTGLPMPKIEWSKNETVIEKPTDALQITKEEVSSEAKTELSIPKAVREDKGT 8977
QY 148 ----- 147
Db 8978 YVTASNRLGVSFRNVHVEYDRPSPRNLAVIDIKAESCYLTWDAPLDNGGSEITHYVI 9037
QY 148 ----- 147
Db 9038 DKRDASRKAWEVNTAVERYGWIWKLIPNGQYEFVRVAVNKYISDECKSKVVIQD 9097
QY 148 ----- 147
Db 9098 PYRLPGPGCKPKVLARTKGSMLVSWTPPLDNGSPITGYWLEKREBGSFYWRSRAPIT 9157
QY 148 ----- 147
Db 9158 KVGLKGVFNVPRLLEGVKYQFRAMAINAAGIGPPSPDPVAGDPIPPPGPPSCPEVK 9217
QY 148 ----- 147
Db 9218 DKTKSSISLWKPAPKADGGSPKIGYIVEMQEEGTTDWRVNEPDKLIITTCCECVPNLKEB 9277
QY 148 ----- 150
Db 9278 RXYFRVKAENAGESEPSDTTGEIPATDIOEPEVFDIGAQCCLVCKAQSIIRPAVI 9337
QY 151 ----- 150
Db 9338 KGRPTPKSWEFDGKAKAMKDVHDIPEDAQLETAENSVIIIPCKRSHTGKYSITAK 9397

QY 151 ----- KGF----- 153
Db 9398 NKAGOKTANCRVKVMDVPGPKDLKVS DITRGSCRLSMKMPDDGGDRIKGYVIEKRTID 9457
QY 154 ----- YPSD- 157
Db 9458 GKAWTKVNPDCGTTTFVVPDLSLSEQYFPRVRAENRFGIGPPVETIQTARTARDPIYPPDP 9517
QY 158 ----- 157
Db 9518 PIKLIKGLITKNTVHLSMKPKNDGGSPVTHYIVECLAWDPTGTKEAWROCNKRDEEL 9577
QY 158 ----- 157
Db 9578 QFTVEDLVEGGYEFVRKAVNAAGVSKPSATVGPCDCQCPMPPSIDLKEFMEVEEGTNV 9637
QY 158 ----- 157
Db 9638 NIVAKIKGVPPPTLTWFKAPKKPDNKEPVLDTYHVNKLVDCTCTLVIPQSRSDTGLY 9697
QY 158 ----- 157
Db 9698 TITAVNNLGTASKEMRLNVLGRPPVGP IKFESVSADQMTLSWPPPKDDGSKITNYVI 9757
QY 158 ----- 157
Db 9758 EKREANRKTWVHSSEPKECTYTIPLLEGHEYVFRIMAONKYGIGEPDSEPTARNLF 9817
QY 158 ----- IAVEWE----- 163
Db 9818 SVPGAPDKPTVSSVTRNSMTVNWEEPEYDGGSPVTGYWLEMKDTSKRWKRVNRDPIKAM 9877
QY 164 ----- 163
Db 9878 TLGVSYKVTGLIEGSDYOFRYVAINAAGVPASLPSDPATARDPIAPGPPPKVDTWTK 9937
QY 164 ----- 163
Db 9938 SSADLEWSPPLKDGSKVTGYIVEYKEBGEKEWKGKQEVRGTKLVVTGLKEGAFYKFR 9997
QY 164 ----- 163
Db 9998 VSAVNIAGIGEPGEVTDVIEMKDRLVSPDLQDASVRDRIVVHAGGVIRIIAYVSGKPPP 10057
QY 164 ----- 163
Db 10058 TVTWMNERTLPQEATIIETTAISSMWIKNCORSHOGVYSLAKNEAGERKKTIIVDULD 10117
QY 164 ----- SNGOPENNY----- 172
Db 10118 VPCPVGTFFLAHNLTNESCKLTWFSPEDDGGSPITNYVIEKRESDRRAWTPVYTYTTRQN 10177
QY 173 ----- 172
Db 10178 ATVGLIOGKAVFPRIAAENSIGMGPFVETSEALVIREPITVPERPEDLEVKEVTKNTWT 10237
QY 173 ----- 172
Db 10238 LTWNPPKYDGGSEIINYVLESRLIGTEKFHKVTNDNLLSRKTYVKGLKEGDTYEYRSVAV 10297
QY 173 ----- 172
Db 10298 NIVGOGKPSFCTKPTCKDELAAPTTLHLDPRDKLTIRVGEAFALTGRYSGRPKPKVSWFK 10357
QY 173 ----- KTFP----- 176
Db 10358 DEADVLEDDRTHIKTTPATLALEKIKAKRSDSGKYCVVVENSTGSRKGFCQVNVVDHGP 10417
QY 177 ----- 176
Db 10418 PVGPVSFDEVTKDMVISMKPPDLDDGSKITNYIIIEKKEVGKDMVMPVTSASAKTTCKVS 10477

QY 177 ----- 176
Db 10478 KLEGGDYIFRIHAENLYGISDPLVSDSMKAKDRFRVPADQPIVTEVTKDSALVTWNK 10537
QY 177 ----- 176
Db 10538 PHDGGKPTITNYILEKRETKSRWARVTKDPIHPTKFRVPLDLEGGQCYEPRVSAENEIGI 10597
QY 177 ----- 176
Db 10598 GDSPPSKPVFAKDPIAKSPVPVNPDAIDTTCNSVDLTWQPPRHDGGSKILGYIVEYQKV 10657
QY 177 ----- 176
Db 10658 GDEWRANHTPESCPETKYKVTGLRDGQTYKFRVLAVNAAGESDPAHVPEPVLVKDRLE 10717
QY 177 ----- 176
Db 10718 PPELILDANMAREQHIVKGDPTLRLSAIIKGVPPFKVTWKKEDRDAPTAKRIDVTPVGSKL 10777
QY 177 ----- 176
Db 10778 EIRNAHEDGGIYSLVNPAGSKTVSVKVLVDKPGPPRDLEVSEIRKDSCHLTWKEPL 10837
QY 177 ----- 176
Db 10838 DDGGSVITNVVRRDVASQAQWSPLSATSKKSHFAKHLNEGNOYLFRVAENQYGRGPF 10897
QY 177 ----- 176
Db 10898 VETPKIKALDPLHPGPKDLHHVDKTEVSLVMNKPDRDGGSPITGYLVEYQEBGTQ 10957
QY 177 ----- 176
Db 10958 DWIKFKVTNLECVTGLQCKYTRFRVKAENIVGLGLDPTTPIECQEKLVPPSELDV 11017
QY 177 ----- 176
Db 11018 KLIEGLVVKAGTVRRPAIIIRGVPTAKWTTDGESEIKTDEHTVETDNFSSVLTIKNCL 11077
QY 177 ----- 176
Db 11078 RRDTEYQITVSAAGSKTVAVHLTVLDVPGPTGPINILDVTPHEMTISWQPPKDDGGS 11137
QY 177 ----- 176
Db 11138 PVINIVEKQDTRKDTWGVVSSGSSKTKLIPHLOKCEYVFRVRAENKIGVGPLDSTP 11197
QY 177 ----- 176
Db 11198 TVAKHKFSPSPGKPVVTDITENAATVSWTLPKSDGSGPITGYIMERREVTGKWRVVK 11257
QY 177 ----- 176
Db 11258 TPIADLKFRVTGLYEGNTYEFVRFAENLAGLSKPSSPDPIKACRPKPPGPPINPKLD 11317
QY 177 ----- 176
Db 11318 KSRETADLVWTKPLSDGSPILGVVVEQCPGTQAWNRIKNDLIRQCAFVPGLIEGNE 11377
QY 177 ----- 176
Db 11378 YRFRIKAANIVGEGEPRELAESVIAKDILHPPEVELDVTCRDVITVRVGOTIRILARVKG 11437
QY 177 ----- 176
Db 11438 RPEPDITWTKGKVLVREKRVLDLIQDLPRVELOIKEAVRADHGKYIISAKNSSCHAQGSA 11497
QY 177 ----- 176
Db 11498 IVNVLDPRPGCONLKVNTVTKENCTISWENPLDNGSGSEITNFIVEYRKNQKGSIVASD 11557
QY 177 ----- 176

Db 11558 VTKRLIKANLANNEYFRVCAENKVGCGPTIETKTPILAINPIDRGPENLHIADKKG 11617
QY 177 ----- 176
Db 11618 TFVYLKWRRPDYDGGSPNLsyHVERRLKGSDDWERVHKSGIKETHYMVDRCVENOIYEFR 11677
QY 177 ----- 180
Db 11678 VQTKNEGSDWVKTBEVVVKEDLOKRPVLDLKLGSVLTVKAGDTIRLEAGVRGKPPPEVA 11737
QY 181 ----- 193
Db 11738 WTKDKDATDLTRSPRVKIDTRADSSKFSKAKRSDGGKYVVTATNTAGSFVAYATVNVL 11797
QY 194 ----- 193
Db 11798 DKPGPVNRLKIVDVSSDRCTVCWDDPDDGGCEIQNYILEKCTKRMVWSTYSATVLTGP 11857
QY 194 ----- 195
Db 11858 TTVTRLIEGNEYIFRVRAENKIGTPTESKPVIAKTKYDKPGPDPEVTKVSKBEMTV 11917
QY 196 ----- 195
Db 11918 VWNPEYDGGKSTGYFLEKKEKHSTWVPVKNKSAIPERRMKYQNLPLDHEYQFRVKAEN 11977
QY 196 ----- 195
Db 11978 EIGEPFSLPSRPVAKDPIEPFGPPTNFRVDTTKHSITLGMGKPYDGGAPIGYVVE 12037
QY 196 ----- 195
Db 12038 MRPKIADASDEGWKRCNAAAQLVRKEFTVTSIDENQYEFYCAQNOVGIGRPAELKEA 12097
QY 196 ----- 195
Db 12098 IKPKEILEPPEIDLDASMRKLIVIRAGCPIRLFAIVGRPAKPVTRKVGIDNVVRKGQV 12157
QY 196 ----- 195
Db 12158 DLVDTMAFLVIPNSTRDDSGKYSLLTVNPAGEKAVFVNVVRVLDTPGPVSDLKVSVDVTKTS 12217
QY 196 ----- 195
Db 12218 CHVSWAPPENDGSGQYTHYIVEKREADRKTWSTVTPVEKTSFHTVNLVPGNEYFRVTA 12277
QY 196 ----- 195
Db 12278 VNEYGCVPTDVPKPVILASDPLSEPDPPRKLATEMTKNSATLAWLPLRDGGAKIDGYI 12337
QY 196 ----- 195
Db 12338 ISYREEQPADRWTEYSVVKDLSLVVTGLKEGKKYKFRVAARNAVGSVLPREAGVYEAK 12397
QY 196 ----- 202
Db 12398 EQLLPKILMPEQITIKAGKKLRIEAHVYKPHPTCKWKKGEDEVVTSSSHVLAHVHKADSSS 12457
QY 203 ----- 202
Db 12458 ILIIKDVTRKDSGYISLTAENSSGTDQKIKVVVMDAPGPPQPFDISIDADACSLSWH 12517
QY 203 ----- 202
Db 12518 IPLEDGGSNTINYIVEKCDVSRGMDWTALASVTKTSCRVKLIPGQBYIFRVRAENRFGI 12577
QY 203 ----- 202
Db 12578 SEPLTSPKVAQPFPGVSEPKNARVTKVNKDCIFVAWDRPDSGSGSPIIGYLIERKERN 12637
QY 203 ----- 202

Db 12638 SLLWVKANDTLVRSTTEYPCAGLVEGLEYSFRITYALNKGAGSPSPKTEYVYTARMVPDPG 12697
QY 203 ----- 202
Db 12698 KPEVIDTKSTVSLIWARPKHDGSKIIGYFVEACKLPCKDKVRCNTAPHQIPOEYAT 12757
QY 203 ----- 202
Db 12758 GLEEKAOYQFRAITARTAVNISPPSPDPVTILAENVPPRIDLSVAMKSLTLVKAGTNCV 12817
QY 203 ----- 202
Db 12818 LDATVFGKPMPTVSMKKGDTLLKPAEGIKMAMQRLCTLEFSVNRKDSGDTITAENS 12877
QY 203 ----- 202
Db 12878 GSKSATIKLVLDKPGPPASVKINKMYSDRAMLSWEPPLDGGSEITNIVDKRETSRN 12937
QY 203 -----VFSCV----- 208
: ||||
Db 12938 WAQVSATVPTITSCSEKLEHGEYQFRIACAENKYGVDVFTEPAIAKNPYDPPGRCDDP 12997
QY 209 ----- 208
Db 12998 VISNITKDHMTYSWKPPADDGSGPITGYLLEKRETOAVNWTKNRKPIERTLKATGLQE 13057
QY 209 ----- 208
Db 13058 GTEYEFVTAINKAGPKPSDASKAAYARDPOYPPAPPAPKVVYDTTRSSVSLSWGKPAY 13117
QY 209 ----- 208
Db 13118 DGGSPILGYLVEBKRADSDNWVRCLPQNLQKTRFEVTGLMEDTOYQFRVAVNKGISD 13177
QY 209 ----- 208
Db 13178 PSDVPDKHYPKDILIPPEGEHDADLRKTLILRAGVTMRLYVPVKRPPPKITWSKPNVL 13237
QY 209 ----- 208
Db 13238 RDRIGLDIKSTDFELRCENVNKYDAGKYILTLNENCGKKEYTIVVKVLDTPGPPINVT 13297
QY 209 ----- 208
Db 13298 VREISKDSAYVTWEPPIIDGGSPILINVVQKRDAERKSWSTVTTECSKTSFRVPNLEBK 13357
QY 209 ----- 208
Db 13358 SYFFRVAENYIGDGPGETRDVAKASQTPGPVVDLKVRSVSKSSCSIGWKKPHSDGGS 13417
QY 209 ----- 208
Db 13418 IIGYVVDLFTENKQVRNKSLSLOYSAKDLTEGKEYTFRVSAENEGEGTPEITVVAR 13477
QY 209 ----- 208
Db 13478 DDWAPDLDLKGLDPLCYLAKENSFRLKIPIKGPAPSVSKKGEDPLATDTRVSVESS 13537
QY 209 ----- 208
Db 13538 AVNTTLIVDQCKSDAGKYTITLKNVAGTKEGTISIKVYGRPIPTGPIKFDEVTAEMT 13597
QY 209 -----MHEALH----- 214
: |||
Db 13598 LKWAPKDDGSEITNILEKRDSDNNKWTVCASAVQKTTFRVTLRHEGMEYTFRVSAN 13657
QY 215 ----- 214
Db 13658 KYGVGELKSEPIVARHPFDVPDAPPPNIVDRHDSVSLWTDPKKTGGSPITGYHLEF 13717
QY 215 ----- 214
Db 13718 KERNSLLWKRANKTPIRMDFKVTGLTEGLEYEFRVMAINLAGVCKPSLPSEPVVLDPI 13777

QY 215 ----- 214
Db 13778 DPPGKPEVINITRNSVTLLIWTPEPKYDGGHKLGTGYIVEKRDLPKSKWMKANHVNPECAFT 13837
QY 215 ----- 214
Db 13838 VTDLVGGKYEPRIRAKNTAGAISAPSESTETIICKDEVEAPTIVLDPTIKDGLTIKAGD 13897
QY 215 ----- 214
Db 13898 TIVLNAISILGKPLKPSKSKAGDIRPSDITQITSTPTSSMLTIKYATRKDAGEYTITA 13957
QY 215 ----- 214
Db 13958 TNPFGTKVEHVKVTVLDVPGPPGPVEISNVSAEKATLTWTTPLEDGGSPIKSYILLEKRET 14017
QY 215 -----NH----- 217
: |||
Db 14018 SRLLTWTVSEDIQSCRHVATKLIQOGNEYIFRVSANVHYKGEPVQSEPVKMWDRFGPPG 14077
QY 218 ----- 217
Db 14078 PEKPEVSNVTNTATVSWKRPVDDGSEITGYHVERREKSLRWVRAIKTPVSDLRCKVT 14137
QY 218 -----TQKSLSLS----- 225
: |||
Db 14138 GLOBEGSTYEFVRSAENRAGIGPPSEASDVLMDAAYPPGPPSNPHVTDTTKKKSASLAWG 14197
QY 226 ----- 225
Db 14198 KPHYDGGLEITGYVVEHQKVGDEAWIKDTTGTALRITQVVPDLQTKKEYNFRISAINDA 14257
QY 226 ----- 225
Db 14258 VGGEPAVIPDVEIVEREMAPDFELDAELRRTLVRAGLSIRIFVPIKGRPAPEVTWTKDN 14317
QY 226 ----- 225
Db 14318 INLKNRANIENSTFTLLIIPECNRYDTGKFWMTIENPAGKSGFVNVRVLDTPGPVNL 14377
QY 226 ----- 225
Db 14378 RPTDITKDSVTLHWDPLIDGSGRITNIVEKREATRKSYSTATTKCHKCTKYVTGLSEG 14437
QY 226 ----- 225
Db 14438 CEYFFRVAENYIGIGETPETTEPVKASEAPSPDLSLIMDITKSTVSLAWPKPKHDGGS 14497
QY 226 ----- 225
Db 14498 KITGYVIEAQRKSGSDQWTHITTVKGLECVVRNLTEGEYTFQVMVNSAGRSAPRESRPV 14557
QY 226 ----- 225
Db 14558 IVKEQTMPLPELDLRGIYQKLVIAKAGDNIKVEIPVLRPKPTVTWKGDQILKQTVRNF 14617
QY 226 ----- 225
Db 14618 ETTATSTILNTNECVSRSDSGPYPLTARNIVEGVDVITIQVHDIPGPTGPIKFDEVSSD 14677
QY 226 ----- 225
Db 14678 FVTFSWDPENDGGVPISNYVVMRQDSTTWVELATTVIRTTYKATRLTGLLEYQFRVK 14737
QY 226 ----- 225
Db 14738 AQNRVGVPGITSAWIVANVPKVPVPGPTQVTAVTAKDSMTISWHEPLSDGSGPILGYH 14797
QY 226 ----- 225
Db 14798 VERKERNILMOTVSKALVPGNIFKSSGLTDGIAIEFRVIAENMAGSKSPKSPSEPMAL 14857

QY	226	-----	225
Db	14858	DPIDPGKPVPLNTRHTVTLKWKAPYTGFGKITSYIVEKRDLPNGRWLKNFSNLN	14917
QY	226	-----	225
Db	14918	EFTVSGLTDAAYEFVRVIAKNAAGAISSPSEPSDAITCRDDVEAPKIKVDVKFKDTVLK	14977
QY	226	-----	225
Db	14978	AGEAFRLADVSRPPTMEWSKDGKELEGTAKLEIKIADFSTNLVKNKDSRRDSCAYTL	15037
QY	226	-----	225
Db	15038	TATNPGGFAKHFNKVLDRPGPEGPLAVTEVTSEKCVLSWFPPLDDGGAKIDHVIQK	15097
QY	226	-----	225
Db	15098	RETSRLAWTNVASEVQVTKLVTKLLKGNEYIFRVMVANKYGVGEPLSEPVLAVNYPGP	15157
QY	226	-----	225
Db	15158	PDPKPNKPVTTITKDSMVVCHGHPDSGSGSEIINIVERRDKAGORWIKCNKKTLDLRY	15217
QY	226	-----	225
Db	15218	KVSGLTEGHEYEFRMAENAGISAPSPSPFYKACDTPFKPGPGNPRVLDTSRSSISI	15277
QY	226	-----	225
Db	15278	AWNKPIYDGGSEITGYMVEIALPEDEWQIVTPPAGLKATSYITITGLTENQEKIRIYAM	15337
QY	226	-----	225
Db	15338	NSEGLGEPALVPCTPKAEDBMLPEIELDADLRKVVTIRACCTLRLEVPKIGRPDPEVKW	15397
QY	226	-----	225
Db	15398	ARDGESLDKASTESASYTLIIVGNVNRFDGKYILTVENSSGSKSAFVNRVLDTPGP	15457
QY	226	-----	225
Db	15458	PQDLKVKVTKTSVTLTWDPPLDGGSKIKNYIVEKRESTRKAYSTVATNCHTKSWKVDQ	15517
QY	226	-----	228
Db	15518	LQEGCSYIFRVAENEGIGLPAETAESYKASERPLPPGKITLMDVTRNSVSLSEKPEH	15577
QY	229	-----	228
Db	15578	DGGSRIILGYIVEMQTKGSKWATCATVKVTEATITGLIOGEEYSFRVSAQNEKGISDPRQ	15637
QY	229	-----	228
Db	15638	LSVPVIAKDLVIPAPKLLFNFTVLAGEDLVDPFPFGRPTPAVTWHKDNVPLKQTRV	15697
QY	229	-----	228
Db	15698	NAESTENSLTITKDACREDVGHVVVKLTNSAGEAIELTNVIVLDKPGPTGPKMDEV	15757
QY	229	-----	228
Db	15758	ADSIITLSWGPYPKYGSSINNYIVEKRDSTTTWQIVSATVARTTIKACRLKTGCEYQFR	15817
QY	229	-----	228
Db	15818	IAAENRYGKSTYLNSEPTVAQYKVPKPGPTPVVTLSSRDSMEVQWNEPISDGSRVIG	15877
QY	229	-----	228
Db	15878	YHLERKERNILWVKNKPIPOTKFTTGLEBGEVEYFRVSAENIVGKPKSKVSECYV	15937
QY	229	-----	228

Db	15938	ARDCDPPGRPEAIIVTRNSVTILQWKKPTYDGGSKITGYIVEKKELPEGRMKASFTNII	15997
QY	229	-----	228
Db	15998	DTHFEVTGLVEDHRYEFRVIARNAAGVSESESTGAITARDEVDPRIISMDPKYKDTIV	16057
QY	229	-----	228
Db	16058	VHAGESFKVDADIYKPIPTIOWIKGOELSNRTARLEIKSTDFATSLSVKDAVRVDSGNY	16117
QY	229	-----	228
Db	16118	ILKAKNVAGERSVTNVKVLDRPPEGPVVISGVTAEKCTLAWKPLQDGGDIINYIV	16177
QY	229	-----	228
Db	16178	ERRETSRLVMTVDANVQTLSCVKTKLLEGNEYTFRIMAVNKYGVGEPLSEPVVAKNPF	16237
QY	229	-----	228
Db	16238	VVPDAPKAPVTTITKDSMIVVWERPASDGGSEILGYVLEKRDKEGIRWTRCHKRLIGEL	16297
QY	229	-----	228
Db	16298	RLRVTGLIENHDYEFVRVSAENAGLSEPPSPAYQACDPIYKPGPNPKVIDITRSSV	16357
QY	229	-----	236
Db	16358	FLWSMKPIYDGGCEIOGYIVEKCDVNVGEWTCTPTPTGINKTNIEVEKLEKHEYNPRIC	16417
QY	237	-----	236
Db	16418	AINKAGVGEHADVPPIIVEEKLAPDIDLLELRKIINIRAGSURLFVPIKGRPTPEV	16477
QY	237	-----	236
Db	16478	KMGKVDGEIRDAAIIDVTSFTSLVLDNVNRYDSGKYTLTLENSSGTKSAFVTVRLDTP	16537
QY	237	-----	236
Db	16538	SPVNLKVTEITKDSVSITWEPPLDGGSKIKNYIVEKREATRKSAAVVTNCHKNSWKI	16597
QY	237	-----	240
Db	16598	DOLQEGCSYIFRYTAENEYIGLPAQTADPIKVAEVPQPPGKITVDDVTNSYSLSWTKP	16657
QY	241	-----	240
Db	16658	EHDGSKIIQYIVEMOAKHSEKSEKARVKSLOAVITNLTOGEEYLFRRVAVANEKGRSDP	16717
QY	241	-----	240
Db	16718	RSIAVPIVAKDLVIEPDKPAFSSYSVQVQDLMKEVPIISGRPKPTITWTKDGLPLKQTT	16777
QY	241	-----	240
Db	16778	RINVTDSDLTLTILSIKETHKDDGGQYITVANVVGOKTASIEIVTLDKPDPKPGVKFDD	16837
QY	241	-----	240
Db	16838	VSAESITLSWNPPLYTGCCQITNVIQKRDTTITVMDVVSATVARTTLKVTCLKTGTEYQ	16897
QY	241	-----	240
Db	16898	FRIFAENRYQSFALESDFIVAQYKYKEPGPGPTPFATAISKDSMVQIHWEPVNNGSPV	16957
QY	241	-----	240
Db	16958	IGYHLERKERNILWTKVKNKTIIDHTQFKAQNLEBEGIEYFRVVAENIVGVGKASNSEC	17017
QY	241	-----	240

Db 17018 YVARPCDPPGTPPEPIMVKRNEITLQWTKPVYDGGSMITGYIVEKRDLPDGRMKASFTN 17077
QY 241 ----- 240
Db 17078 VIETQTVSGLTEDQRYEPRVIAKNAAGSAIKSPSDSTGPITAKDEVLPRIKMDPKFRDT 17137
QY 241 ----- 240
Db 17138 IVVNAGETFRLEADVHGRKPLPTIENLGRDKIEESARCEIKNTDFKALLIYKDAIRIDGG 17197
QY 241 ----- 240
Db 17198 QYILRASNVAGSKSPVNVKVLDRPPEGPVQVGTGVTSEKCLTWSPPLODGGSDISHY 17257
QY 241 ----- 240
Db 17258 VVEKRETSRLAWTVVASEVVTNSLKVTKLLEGNVYFRIMAVNKYGVGEPLSAPVLKMN 17317
QY 241 ----- 240
Db 17318 PFVLPGPKSLEVTNIANDSMIVCHNRDSDGSEIIGYIVEKRDORSIGIRWIKCNKRIT 17377
QY 241 ----- 240
Db 17378 DLRLRVTLGTEDHEYEFVSAENAGVGEPSPATVYKACDPVKPGPPTNAHIVDTTKN 17437
QY 241 ----- 240
Db 17438 SITLAWKPIYDGGSEIIGYVEICKADEEWOIVTPOTGLRVTRFEISKLTEHOEYKIR 17497
QY 241 ----- 240
Db 17498 VCALNKVGLGEATSVPGTVKPEDKLEAPELDDSELKRGIVVRAGGSARIHIPFKGRPMP 17557
QY 241 ----- 240
Db 17558 EITWSREEGEFTKVQIEKGVNYTOLSIDNCDNRDAGKYILKLENSGSKSAFVTVKVL 17617
QY 241 ----- 240
Db 17618 TPGPPONLAVKVRKDSAFVWEPPIIDGGAKVKNYVIDKRESTRKAYANVSSKCSKTSF 17677
QY 241 ----- 240
Db 17678 KVENLTEGAIYFRVMAENEFVGVPVETVDAVKAAPPPGPKYTLTDVDSQTSASLWME 17737
QY 241 ----- 240
Db 17738 KPEHGGSRVLGYVVMQPKTEKWSIVAESKVCNAVVTGLSSGOEYQFRVKAYNEKGS 17797
QY 241 ----- 240
Db 17798 DPRVLGVPVIAKDLTIQPSLKLPIFNFTYSIQAGEDLKIPIVIGRPRPNISWVKDEPLKQ 17857
QY 241 ----- 240
Db 17858 TTRVNVEETATSTVLHIKEGNKDDPGKYTVTATNSAGTATENLSVIVLEKPGPPVGPVRF 17917
QY 241 ----- 240
Db 17918 DEVSADFVVISWEPPAYTGGCOISNYIVEKRDTTTTTHWVSATVARTTIKITLKTGTE 17977
QY 241 ----- 240
Db 17978 YQFRIFAENRYKSAPLDSKAVIVQYPFKEPGPGTPTVTSISKDQMLVQWHEPVNDGGT 18037
QY 241 ----- 240
Db 18038 KIIGYHLEQKEKNSILWVKNLKIPIQDTKFKTGTLDEGLEVEFKVSAENIVGIGKSKVS 18097
QY 241 ----- 240
Db 18098 ECFVARDPCDPPRPEAIVITRNNVTWKKKPAYDGGSKITGYIVEKRDLPDGRMKASF 18157

QY 241 ----- 240
Db 18158 TNVLETEFTVSGLVEDQRYEPRVIARNAAGNFSEPSDSSGALTARDEIDAPNASLDPYK 18217
QY 241 ----- 240
Db 18218 DVIVVHAGETFVLEADIRGKPIPDVVVMSKDGKLEETAARMEIKSTIQKTTLLVWKDCIRT 18277
QY 241 ----- 240
Db 18278 DGGQYILKLSNVGTKSIPITVKVLDRPGSPGPKLVGTVAEKCYLAWNPPLODGGANI 18337
QY 241 ----- 240
Db 18338 SHYIEKRETSRLSWTQVSTEQVALNWKTKLLPGNEYIFRVMVKNKYGIGEPLESGPVT 18397
QY 241 ----- 240
Db 18398 ACNPYKPPGPPSTPEVSAITKDSMVVVTWARPVDDGGTEIEGYILLEKRDKEGVRWTKCNKK 18457
QY 241 ----- 240
Db 18458 TLTDLRLRVTLGTGHSYEFVRAAENAGVGEPSSEVYRACDAILYPPGPSNPKVTD 18517
QY 241 ----- 244
Db 18518 SRSSVSLAWSKPIYDGGAPVKGYVVEKAAADEWTTCTPPTGLQGKQFTVTKLENTEY 18577
QY 245 ----- 244
Db 18578 NFRICAINSEGVGEPATLPGSVVAQERIEPPEIELDADLRKVVVLRSATLRLFTIKGR 18637
QY 245 ----- 244
Db 18638 PEPEVKWEKABGILTDRAQIEVTSFTMLVIDNVTRFDSGRYNTLENNSGSKTAFNVNR 18697
QY 245 ----- 244
Db 18698 VLDSPAPVNLTIREVKKDSVTLSEPPPLIDGGAKITNYIVEKRETTTRKAYATITNNCTK 18757
QY 245 ----- 244
Db 18758 TTFRIENLOEGCSYFRVLASNEYGIGLPAETTEBPVKVSEPPLPBGRVTLVDVTRNTATI 18817
QY 245 ----- 244
Db 18818 KWKEPESDGGSKITGYVVMQTKGSEKWSCTQVKTLEATISGLTAGEEYVFRVAANEK 18877
QY 245 ----- 244
Db 18878 GRSDPRQLGVPVIAARDIEIKPSVELPHTFNVKAREQLKIDVPFKGRPQATVNNRKDQOT 18937
QY 245 ----- 244
Db 18938 LKETTRVNVSSKTVTSLSIKEASKEDVGYELCVNSAGSITVPTIILVLDLRPGPGPI 18997
QY 245 ----- 244
Db 18998 RIDEVSCDSITISWNPPEYDGGCOISNYIVEKETTSTTHWVSQAVARTSIKIVRLTTG 19057
QY 245 ----- 244
Db 19058 SEYOFRVCAENRYCKSSYSESSAVVAEYFSPPGPGTPKVVHATKSTMLVTHQVPVNDG 19117
QY 245 ----- 244
Db 19118 GSRVIGYHLEYKERSILLWSKANKILIADTQVKVSGLDEGLMEYRVVAENIAGIGKSK 19177
QY 245 ----- 244
Db 19178 SCEVPARDPCDPPGQPEVTNITRKSLSKWSKPHYDGGAKITGYIVERRELDPGRMLKC 19237

QY	245	-----	244
Db	19238	NYTNIQETVEFTELTEDQRYEPRVFARNAADSVSEPSBSTGPIIVKDDVEPRVMDVK	19297
QY	245	-----	244
Db	19298	FRDVIIVKAGEVLKINADIAGRLPVISWAKOGIEIERARTEIISTDNHLLTVKDCIR	19357
QY	245	-----	244
Db	19358	RDFGQYVLLKNVAGTRSAVNCKVLDKPGPPAGPLEINGLTAEKCSLSWGRQEDGGAD	19417
QY	245	-----	244
Db	19418	IDYHRRKKRETSLAWTICEGELQMTCKVTLLKNGEYIFRVTGVNKYGVGEPLSVAI	19477
QY	245	-----	244
Db	19478	KALDPFTVPSPPTSLEITSVTKESMTLCWSRPESDGSSEISGYIIBERREKNSLRWRVVK	19537
QY	245	-----	244
Db	19538	KPYDLRVKSTGLREGCEYEYRYVAENAGLSLPSETSPLIARAEDPVFLPSPSPKPIVD	19597
QY	245	-----	244
Db	19598	SGKTTITIAWVKPLFDGAPITGYTVEYKKSDDTDKWTSIQSLRGTEYTIISGLTTGAEVY	19657
QY	245	-----	244
Db	19658	FRVKSVMKVGASDPSSDPQIAKEREEREPFLDIDSEMRTKLIVKAGASFTMTVPFRGR	19717
QY	245	-----	244
Db	19718	VPNVLSKPDTLTRAYVDTDSRTSLTIENANRDSGKYTLTIQNVLSAASLTIVKV	19777
QY	245	-----	244
Db	19778	LDTPGPTNITVQDVTKEAVLSWDVPENDGGAPVKNYHIEKREASKAWSVTNCNRL	19837
QY	245	-----	244
Db	19838	SYKVTNLQEGAIYFRVSGENEFVGIPAEATKEGVKITEKPSPEKLGVTISIKDSVSLT	19897
QY	245	-----	244
Db	19898	WLKPEHGGSRIVHYVVEALEKGQKNWKCATAKSTHHVVSGLRENSEYFFRVAENQAG	19957
QY	245	-----	244
Db	19958	LSDPRELLLPVLKEQLEPPEIDMKNFPSHTVYVRAGSNLKVDPISGKPLPKVTLSDG	20017
QY	245	-----	244
Db	20018	VPLKATMRENTETAENLTINLKESVTADAGRYEITAANSSGTTKAFINIVLDRPGPPT	20077
QY	245	-----	244
Db	20078	GPVVISDITEESVTLKWEPPKYDGGSOVTNYILLKRETSTAVWTEVSATVARTMKVMKL	20137
QY	245	-----	244
Db	20138	TTGEEYQFRIKAENREGISDHIDSACVTVKLPYTPPGPSTPWNTVRESITVGNHEPV	20197
QY	245	-----	244
Db	20198	SNGGSAAVGYHLEMKDRNSILMQANKLVIRTHFKVTTISAGLIYEFRRYAENAAVGK	20257
QY	245	-----	247
Db	20258	PSHPSEPVLAIDACEPRNVRITDISKNSVLSWQQPAFDGSGKITGYIVERRDLPDGRW	20317
QY	248	-----	247
Db	20318	TKASFNTVNTQFTTISGLTQNSQYEFVRFARNAVGSISNPSEVYVGPITCIDSYGGPVIDL	20377
QY	248	-----	247
Db	20378	PLEYTEVVKYRAGTSVKLRAGISGKPAPTIETWKDDKDELQTNALVCVENTTDLASILKD	20437
QY	248	-----	247
Db	20438	ADRLNSGCVELKLRNAMASASATIRVQILDKPGPGPIEFKVTVAEKITFLWRPPADDG	20497
QY	248	-----	247
Db	20498	GAKITHYIIVEKRETSRWVMSWSEHLEECIITTKIIGNEYIFRVRANVKYIGIGEPLES	20557
QY	248	-----	247
Db	20558	DSVVAKNAFVTPGPGPIPEVTKITKNSMTVWNSRPIADGSDISGYFLEKRDKSLGWFK	20617
QY	248	-----	247
Db	20618	VLKETIRDTRQKVTGLTENSQYRVCANNAAGOGPSESEFYKAADPIDPPGPPAKIR	20677
QY	248	-----	247
Db	20678	IADSTKSSITLWGSKPYDGSVATGYVVEIROGEEEEWTVTSTKGEVRTTEYVWSNLKP	20737
QY	248	-----	247
Db	20738	GVNYIFRVSANVCAGOGEPNEMNEPVQAKDILEAPEIDLDVALRTSVIAKAGEDVQLIP	20797
QY	248	-----	247
Db	20798	FKGRPPPTVTRKDEKNLGSARYSIENTDSSILLTIPOVTRNDTGKYILTIENGVEPK	20857
QY	248	-----	247
Db	20858	SSTVSVKVLDTPAACOKLOVKHVSRTVTLLWDPPLIDGGSPINIVYIEKRDTAKRTWSV	20917
QY	248	-----	247
Db	20918	VSHKCSSTSFKLIDLSEKTPFFFRVLAENEIGICEPCETTEPVKAAEVPAPIRDLMSKDS	20977
QY	248	-----	247
Db	20978	TKTSVILSMTKPDFDGGSVITEYVVERKKGEGQWTHAGISKTCEIEVSQLEKQSVLEFR	21037
QY	248	-----	247
Db	21038	VFAKNEKGLSDPVTIGPITVKELIITPEVDLSDIPGAQVTRIGHNVHLELPYKGRPKPS	21097
QY	248	-----	247
Db	21098	ISWLKDGPLKESEFVRFSKTENKITLSIKNAKKEGGKTYTVIDNACVIAPIVITL	21157
QY	248	-----	247
Db	21158	GPSPKPGPIRDEIKADSVILSWDVPEDNGGGEITCYSIEKRETSQTNKMKVCSVART	21217
QY	248	-----	247
Db	21218	TFKVPNLVKDAEYQFRVRAENRYGVSOPLVSSIIIVAKHQFPIPGPKPVIYNVTS DGMS	21277
QY	248	-----	247
Db	21278	LTWDAVPYDGGSEVTGFHVEKKERNILMQKVNTSPISGREYRATGLVEGLDYQFRVYAE	21337
QY	248	-----	247
Db	21338	NSAGLSSPSDPKFTLAVSPVDPPTDYIDVTREITLKNWPNPLDRDGGSKIYGVYSIEKR	21397
QY	248	-----	247

Db 21398 QGNERWVRNFTDVSECOVTVTGLSPGDRYEFRIIARNAVGTISPPSQSGIIMTRDENV 21457
QY 248 ----- 247
Db 21458 PPIVERGPEYFDGLIIKSGESURIKALVOGRVPRVTFKDGVEIEKRMNMEITNVLGST 21517
QY 248 ----- 247
Db 21518 SLFVRDATRDHRGVYTVAKNAGSAAKAEIKVKVQDTPGKVVGPIRFTNITGKMTLWWD 21577
QY 248 ----- 247
Db 21578 APLNDGAPITHYIIEKRETSRLAWALIEDKCEAQSYTAIKLINGNEYQFRVSAVNKFGV 21637
QY 248 ----- 247
Db 21638 GRPLSDPVAQIOYTVDPAGPIPEPSNITGNSITLTWARPESDGGSEIQOYILERREKK 21697
QY 248 ----- 247
Db 21698 STRWVKVISKRPISETFRFKVTKGLTEGNEYEFHVMAENAAAGVGPASGISRLIKCREPVNPP 21757
QY 248 ----- 247
Db 21758 GPPTVVVKVTDTSKTVSLEWSKPVFDGGMEEIIGYIIECMCKTDLGDWHKVNABEACVKTRYT 21817
QY 248 ----- 247
Db 21818 VTDLQAGEYKFRVSAINGAGKDCSEVTGTTIKAVDRLTAPELDIDANFKQTHVVRAGAS 21877
QY 248 ----- 247
Db 21878 IRLFIAYOGRPTTAVWSKPDNSLSRADIHFTDSFSTLTVCNCRNDAGKYTLAVENNS 21937
QY 248 ----- 247
Db 21938 GSKSITFTVKVLDTPGPPGPIITFKDVRGSAFLMWDAPLLDGGARIHHYVVEKREASRRS 21997
QY 248 ----- 247
Db 21998 WQVISEKCTRQIFKVNDLAEGVPYFRVSAVNEYGVGPEYEMPEIVATEQPAPRRLDV 22057
QY 248 ----- 247
Db 22058 VDTSKSSAVLAWLKPDPHDGSRITCYLLEMROKGSDLWVEAGHTKOLFTTVERLVEKTEY 22117
QY 248 ----- 247
Db 22118 EFRVAKNDAGYSEPREAFSSVIIKEPQIEPTADLTGITNOLITCKAGSPFTIDVPISGR 22177
QY 248 ----- 247
Db 22178 PAPKVTWKEENRLKETDRVSITTTKDRITLTVKDSMRGDSGRYFLTLENTAGVKTFSTV 22237
QY 248 ----- 247
Db 22238 VVVIGRPGVPTGPIEVSSVSAESCVLWGPCKDGGTEITNYIVEKRESGITAWQLVNSS- 22297
QY 248 ----- 247
Db 22298 VKRTQIKVTHLTKYMEYSFRVSSSENRFVSKPLESAPIAIEHPFVPPSAPTRPEVYHVSA 22357
QY 248 ----- 247
Db 22358 NAMSIRWEEPYHDGSKIIGYVVEKKERTILWVKENKVPCLCNYKVTGLVEGLEYOFR 22417
QY 248 ----- 247
Db 22418 TYALNAGYSKASEASRPIMAQNPVDPAGRPEVTDVTRSTVLSIWSAPAYDGGSKVGYI 22477
QY 248 ----- 247
Db 22478 IERKPVSEVGDRWLKCNVTIVSDNFTVTALSEBDTYEFRVLAKNAAGVISKSESTGP 22537

QY 248 ----- 247
Db 22538 VTCRDEYAPPKAELDARLHGLVITRAGSDLVLDAAVGGKPEPKIITWTKGKELDCEKV 22597
QY 248 ----- 247
Db 22598 SLOYTGKRATAVIKFCDRSDSGKYTLTVKNASGTRKAVSMVKVLDSPGCGKLTYSRVTO 22657
QY 248 ----- 247
Db 22658 EKCTLAWSLPQBDGAEITHYIVERRETSRLNMWIVEGECPTLSYVVVTRLIKNNEYIFRV 22717
QY 248 ----- 247
Db 22718 RAVNKYGPVPESEPIVARNSTFTIPSPGPIEEVGTGKEHIIIOWTKPESDGGNEISNY 22777
QY 248 ----- 247
Db 22778 LVDRKREKSLRWTRVYNKDYVVYDTRLKVTSLMEGCDYQFRVTAVNAAGNSEPSESRNFTS 22837
QY 248 -----WLKAFYD----- 254
Db 22838 CREPSYTFGPPSAPRVVDTTKHSISLAWTKPMYDGGTDIVGVLEMQEKTDQWYRVHTN 22897
QY 255 ----- 254
Db 22898 ATIRNTEFTVPLKMGQKYSFRVAAVNVKGMSEYSESTAEIEPVERIERIPDLELADLLK 22957
QY 255 ----- 254
Db 22958 TWTIRAGASRLWVSVSGRPPVITWSKQIDLASRAIIDTTESYSLLIIVDKVNRDAGK 23017
QY 255 ----- 254
Db 23018 YTIEAENOSGKKSATVLVKVYDTPGPCSVKVEKVSRSVTTITWEIPTDGGAPINNYIV 23077
QY 255 ----- 254
Db 23078 EKREAAARAFKTVTKCKSLYRISGLVEGTMHYPRVLPENIYGIGECETSDAVLVSEV 23137
QY 255 ----- 254
Db 23138 PLVPAKLEVDVDTKSTVTLAWEKPLYDGGSLTGYVLEACKAGTRWMKVVTLKPTVLEH 23197
QY 255 ----- 254
Db 23198 TVTSLNEGEQYLFRITRAONEKGVSEPRETVTAVTVQDLRVLPTIDLSTMPQKTIHVPAGR 23257
QY 255 ----- 254
Db 23258 PVELVIPIAGRPPPAASHFFAGSKLRESERVTVETHTKVAKLITRETTIRDTGTYTLEK 23317
QY 255 ----- 254
Db 23318 NVTGTTSETIKVIIILDKPGPTPIKIDEIDATISITISWEPPELDGGAPLSGYVVEQRDA 23377
QY 255 ----- 254
Db 23378 HRPGLPVSESVTRSTFKFTRLTEGNEYVFRVAATNREFGYSYLOSEVIECRSSIRIPGP 23437
QY 255 ----- 254
Db 23438 PETLIQIFDVSRDGMTLTWYPPEDDGGSOVTGYIVERKEVRADRWVRVNVKVPVMTTRYST 23497
QY 255 ----- 254
Db 23498 GLTEGLEYEHRVTAINARGSGKSRPSKPIVAMDP IAPPKGPQNPRVTDTRTTSVSLAWS 23557
QY 255 ----- 254
Db 23558 VPEDGGSKVTGYLIEQKQVDOHEWTKONTPTKIREYTLTHLPOGAEYRFRVLACNAGG 23617

```
QY 255 -----KVAEKL----- 260
|| ||
Db 23618 PGEAEVPGTVKVTMELEYDPYELDERYQEGIFVRQGVIRLTIPKGPCKWTKEG 23677
QY 261 ----- 260
Db 23678 QDISKRAMIATSETHTELVIKEADRGSGTVDLVLENNCKGKAVYIKVRVIGSPNPEGP 23737
QY 261 ----- 260
Db 23738 LEYDDIQVRVSWRPPADGGADILGYILERREVPKAAWTTIDSRVGTSLVVKGLKE 23797
QY 261 ----- 260
Db 23798 NVEYHFRVSAENQFGISKPLKSEEPVTKPLNPPEPPNPPELDVTKSSVLSWSRPK 23857
QY 261 ----- 260
Db 23858 DGGSRVTGYIERKETSTDKVVRHNKTQITTTMYTVGLVPDAEYQFRIIAQNDVGLSE 23917
QY 261 ----- 260
Db 23918 TSPASEPVVKDPDKPSQPGCELEILSISKDSVTLOWEKPCECDGKELGYWVEYRQSGD 23977
QY 261 ----- 260
Db 23978 SAKSKNERIKDKOFTIGGLEATEYEFVFAENETGLSPRRTAMSIKTKLTSGEAPG 24037
QY 261 ----- 260
Db 24038 IRKEMKDVTKLGBAQLSCOIVGRPLPDIKWYRFGKELIQSRKYMSSDGRTHTLVMT 24097
QY 261 ----- 260
Db 24098 EEQDEGYVTCIATNEVGEVETSSKLLLOATPQHPGYPKKEYYGAVGSTLRHLHVMVIG 24157
QY 261 ----- 260
Db 24158 RVPAMTWFGKLLQNSENITIENTHYTHLVKMNQVRKTHAGKYKQLSNVFGTVDAI 24217
QY 261 ----- 260
Db 24218 LDVEIQDKPDKPTGPIVIEALLKNSAVISWKPPADGGSWITNYVVEKKAEGAEQWLV 24277
QY 261 ----- 260
Db 24278 SSAISVTCRIVNLTENAGYFRVSAQNTFGISDPLEVSSVVIKSPFEKPCGPKPTIT 24337
QY 261 ----- 260
Db 24338 AVTKDSCVWANKPPASDGGAKIRNYLEKREKKONKWSVTTEIRETVFSVKNLIEGLE 24397
QY 261 ----- 260
Db 24398 YEFVKCNLGESEWSEISEPITPKSDVP IQAPHFKELNLRVYQSNATLVCKVTGH 24457
QY 261 ----- 260
Db 24458 PKPIVKWYRQKELIADGLKYRIOEFKGYHQLIIASVTDDATVYQVRATNGSVSGT 24517
QY 261 ----- 260
Db 24518 ASLEVEVPAKTHLPKLTLEGMGAVHALRGEVVSIKIPFSGKPDVITWQGDLDLNNGHY 24577
QY 261 ----- 260
Db 24578 QVIVTRFSTSLVPFNGVERKPDAGYVVCANKNRFGIDQKTVELDVADVPDPGRVKVSDAS 24637
QY 261 ----- 260
Db 24638 RDSVNLWTPEASDGGSKITNYIVEKCATTAERWLVRQARETRYVINLFGKTSYQPRV 24697
QY 261 ----- 260
```

```
Db 24698 IAKNFGLSKSPSEPTITTKEDKTRAMNYDEVDETREVSMTKASHSTKELYEKYMA 24757
QY 261 ----- 260
Db 24758 EDLGRGEGIVHRCVETSSKTTYMAKFKVKGTDQVLVKKEISILNIARHRNHLHESF 24817
QY 261 ----- 260
Db 24818 ESMEELWIFFISGLDIFERINTSAFELNEREIVSYHVQCEALQFLHSHNIGHFDIRP 24877
QY 261 ----- 260
Db 24878 ENIIYQTRRSSTIKIIEFGARQLKPGDNFRLLFTAPEYAYEVEHQHDVWSTATDWSLG 24937
QY 261 -----KEAF 264
:||||
Db 24938 TLVVYLLSGINPFLAETNQIIEINMNAEYTFDEEAF 24974

RESULT 15
Q8WZB3
ID Q8WZB3 PRELIMINARY; PRT: 26926 AA.
AC Q8WZB3;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE N2B-titin isoform.
GN TTN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20309627; Pubmed=10850961;
RA Freiburg A., Trombitas K., Hell W., Cazorla O., Fougereousse F.,
RA Centner T., Kolmerer B., Witt C., Beckmann J.S., Gregorio C.C.,
RA Granzier H., Labeit S.;
RT "Series of exon-skipping events in the elastic spring region of titin
as the structural basis for myofibrillar elastic diversity.";
RL Circ. Res. 86:1114-1121(2000).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21573839; Pubmed=11717165;
RA Bang M.L., Centner T., Fornoff F., Geach A.J., Gotthardt M.,
RA McNabb M., Witt C.C., Labeit D., Gregorio C.C., Granzier H.,
RA Labeit S.;
RT "The complete gene sequence of titin, expression of an unusual -700
kDa titin isoform and its interaction with obscurin identify a novel
Z-line to I-band linking system.";
RL Circ. Res. 89:1065-1072(2001).
DR EMBL; AJ277892; CAD12455.1; -.
DR InterPro; IPR000282; Cytok_receptor_2.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR000577; FGGY_kin.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR002016; Peroxidase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF000041; fn3; 132.
DR Pfam; PF000047; ig; 91.
DR Pfam; PF000069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00060; FN3; 132.
DR SMART; SM000409; IG; 111.
DR SMART; SM000408; IGC2; 93.
DR SMART; SM000410; IG_like; 3.
```


QY 43 -----VVVDVSHED----- 51
Db 2078 IYWWPEDNVCLEVRDVTAEASIMVKAINIAGETSSHAFLVQAQKLLITQELQDV 2137
QY 52 -----PEVKFNWYDGVVH----- 66
Db 2138 VAKEKDTMATFECETSEPFVKVWKYKDMGVHEGDKYRMHSDRKVHFLSILITDTSDAED 2197
QY 67 -----NAKT----- 70
Db 2198 YSCVLVEDENKVTAKLIVEGAVVEFVKELQDIEVPSYSGELECI VSPENIEGKWYHND 2257
QY 71 -----KPR----- 73
Db 2258 VELKSNGKYTITSRRGRQNLTVKDVTKEDQGEYSFVIDGKTKCKLMKRPPIAILQGLS 2317
QY 74 ----- 73
Db 2318 DQKVCESDIOLEVKVSVLESVEGVMKMGQEVQPSDRVHIVIDKQSHMLLIEDMTKEDAG 2377
QY 74 ----- 73
Db 2378 NYSFTIPALGLSTSGRSVSVSDVITPLKDVNVIEGTKAVLECKVSPDVTSVKWLINDE 2437
QY 74 ----- 73
Db 2438 QIKPDDRQVAIVKGTOKORLVINRTHASDEGPYKLIIVGRVETNCNLSVEKIKIIRGLRDLT 2497
QY 74 ----- 73
Db 2498 CTETQNVFEVLSHGIDVLNMFNFKDKEIKPSSKYKIEAHGKIYKLTVLNMMKDEGKYT 2557
QY 74 ----- 73
Db 2558 FYAGENITSGKLTVAGAIKSKPLTDQTVABEQAEVPECEVANPDSKGWLRDGHPLTN 2617
QY 74 ----- 73
Db 2618 NIRSESDGHKRLIIAATKLDIDIGEYTKVATSKTSAKLKVEAVKIKTKLNLVTETQD 2677
QY 74 ----- 73
Db 2678 AVFTVELTHPNVGVQWIKNGVNVLESNEYAISVKGTIYSLRIKNCIAIVDESUYGFRLGR 2737
QY 74 ----- 73
Db 2738 LGASARLHVETVKI IKKPKDVTALENATVAFESVSHDTPVVKWFHKNVEIKPSDKRLV 2797
QY 74 ----- 73
Db 2798 SERKVHKLMLQNLSPSDAGEYTA VVGQLECKAKLFVETLHITKTMKNIEVPETKTASFEC 2857
QY 74 ----- 73
Db 2858 EVSHFNVPMSMLKNGVEIEMSEKFI VVOGKHLQLIIMNTSTEDSAEYTFVCGNDQVSAT 2917
QY 74 ----- 73
Db 2918 LTVTPIMITSMKLDINAEKDTITFEVTNYEGISYKWLKNGVEIKSTDKCMQRTKKLTH 2977
QY 74 ----- 73
Db 2978 SLNIRNVHFGDAADYTFVAGKATSTATLYVEARHIEFRKHIDIKVLEKKRAMFECEVSE 3037
QY 74 ----- 73
Db 3038 PDITVOMKDDQELQITDRIKIOKEYVHRLIPLIPSTRMSDAGKYTVVAGNVSTAKLFVE 3097
QY 74 ----- 73
Db 3098 GRDVRIRSIKKEVOVIEKQRAVVEFVNEDDVDHMKDGIENFQVQERHKYVVERRIH 3157
QY 74 ----- 73

Db 3158 RMFISETRQSDAGEYTFVAGNRSSVTLTYVNAPEPPQVLOELQVTVYOSGKPARFCAVIS 3217
QY 74 -----EE----- 75
Db 3218 GRPQPKISWYKEBQLLSTGFKCKFLHDGQBYTLLIIAFAFPDAAVYTCANKNDYGVATTS 3277
QY 76 ----- 75
Db 3278 ASLSVEVPEVVSDQEMPVYPPAIITPLQDVTVISEGQPARFQCRVSGTDLKVSWYSKDKK 3337
QY 76 -----QYNSTYRV----- 83
Db 3338 IKPSRFRMTQFEDTYQLEIAEAYPEDEGTVTYFVASNAGQVSVSTANLSLEAPESILHER 3397
QY 84 ----- 83
Db 3398 IOEOIEMEMKFSFSSFLSABEGLHSAELQLSKINETLELLSESPVYSTKFDSEKGTGP 3457
QY 84 -----VSVL----- 87
Db 3458 IFIKEYSNADISMGDVATLSVTYIGIPKIQHFFNGVLLTPSADYKFEVFDGDDHSLIIL 3517
QY 88 ----- 87
Db 3518 FTKLEDEGEYTCMASNDYGTICISAYLKINSKGEGHKDTETESAVAKSLEKLGCPPHF 3577
QY 88 -----TVLH----- 91
Db 3578 LKELKPIRCAOGLPAIFEYTVWGEPAFTVTFWKENKOLCTSVYTYIIHNPNGSTFIVND 3637
QY 92 ----- 91
Db 3638 POREDSGLYICKAENMLGESTCAEALLVILLEDTDMTDTPCAKSTPEAPEDFPOTPLKGP 3697
QY 92 ----- 91
Db 3698 AVEALDSEQEIATFVKDTILKAALITEENQOLSIEHIAKANELSSQLPLGAQELQSILEQ 3757
QY 92 ----- 91
Db 3758 DKLTPESTREFLCINGSIHFOPLKEPSPNLQLOIVOSQKTFSEKGIILMPEPEPTQAVLSD 3817
QY 92 ----- 91
Db 3818 TEXIFPSAMSIQINSITVEPLKTLAEPEGNYPQSSIEPPMHSYLTSVAEVLSPKEKT 3877
QY 92 ----- 91
Db 3878 VSDTNREQRTLOKQEAQSALILSQSLAEGHVESLQSPVMI SQVNYEPLVPSEHSCTEG 3937
QY 92 ----- 91
Db 3938 GKILIESANPLENAGQDSAVRIEGBKSLRPLALEEKQVLLKEEHSNVVMPDQIIIESK 3997
QY 92 -----QDWL- 95
Db 3998 REPVAIKVQVOGRDILLSKESLLSGIPEQRNLKTIQICRALQAASBQPLFSEWLR 4057
QY 96 ----- 95
Db 4058 NIEKVEAVNITQEPHRHIMCVLTVSAKSVTEEVTTIIEDVDPMANLKMELDALCAI 4117
QY 96 ----- 95
Db 4118 IYEEIDILTAEPRIQOGAKTSLQEMDSFGSOKVEPIPEPEVESKYLISPEEVSYFNV 4177
QY 96 ----- 95
Db 4178 QSRVKYLDATPVTKGVASAVVDEKQDES LKPSKEESESSECTEEVATVKIQEABGGF 4237
QY 96 ----- 95

Db 4238 IKEDGPMIHTPLVDTVSEGDIVHLLTTSITNAKEVNWYFENKLVPSDEKFKCLQDQNTYT 4297
QY 96 -----NGK----- 98
Db 4298 LVIDKVNTEHDHQEYVCEALNDSGKTATSAKLTIVVKRAAPVIRKIEPLEVALGHLAKFT 4357
QY 99 -----EYKCKVSN----- 106
Db 4358 CEIQSAPNVRFOFKAGREIYESDKSIRSKYISLSLEILRTQVVDCGEYTKCASNEYS 4417
QY 107 ----- 106
Db 4418 VSCATLTVTVPGEKKVKRLLPERKPEKEVVLKSVLRKRPEBEEPKVEPKLEKVKK 4477
QY 107 ----- 106
Db 4478 PAVPEPPPKPVEEYVPTVTKRERKIPETKVPKPAIPALPAPKPKPAEAVKTIKP 4537
QY 107 -----KALPAPIE----- 114
Db 4538 PPVEPEPTPIAAPVTVPVVGKKAEPKAAKPKGPIKGVPKKTPSPIEAERRKLRPG 4597
QY 115 ----- 114
Db 4598 SGGEKPPDEAPFYOLKAVPLKVEKIEKIDILTESEFVGSSAIFECVSPSTAITWMKD 4657
QY 115 ----- 114
Db 4658 GSNIRESPKHFADGDKRKLHIIDVQLSDAGEYTCVLRGLNKEKTSTAKLVVEELPVRF 4717
QY 115 -KTIS- 118
Db 4718 VRTLEBEVTVKGPQLYLSCELNERDVVVRKDGKIWEKPGRIVGPVIGLMALTIINDA 4777
QY 119 ----- 118
Db 4778 DTDAGTYTIVTENANNLECCSCVKEVIRDMVLKPIRDQHVKPKGTAIFACDIADKTP 4837
QY 119 ----- 118
Db 4838 NIKWFGYDEIPAENDKTEILRDGNHLYLKIKNAMPEDIAEYAVEIEGKRYPAKLTIGE 4897
QY 119 ----- 118
Db 4898 REVELLKPIEDVTIYEKESASFDAETSEADIPQWKLKCELLRPSPTCEIKAEKGKREF 4957
QY 119 ----- 118
Db 4958 LRKVKLDQAGEVLYOALNAITTAITLVKEIFELDFAVPLKDVTPVRRQARFECVLTREAN 5017
QY 119 ----- 118
Db 5018 VIWSKGPDIKSSDFDIADGKKHILVINDSQFDDGEVYTAIEVEGKTSARLFTVGTGRL 5077
QY 119 ----- 118
Db 5078 KPMSPLEQTVKEGETATFVCELSHEKMHVWFKNDAKLHTSRTLISSEKGTHELMKE 5137
QY 119 ----- 118
Db 5138 VTLDISQIAQVKELSSLAQLKVLADPYFTVKLHDKTAVEKDEITLKCEVSKDVPVKW 5197
QY 119 ----- 118
Db 5198 FKDGEEIVSPKYSIKADGLRLIKKADLKDKGEYVDCGTDKTKANVTVEARLIKVE 5257
QY 119 -----KAGQP----- 124
Db 5258 KPLYGEVFGVGETAHEIELSEPDVHGOWKLKGPLTASPDCEIIEDGKKHILLHNCQL 5317
QY 125 -----REPO----- 128
Db 5318 GMTGEVSQAANAKSAANLKVKEPLIFITPLSDVKVFEKDEAKPECEVSREPKTFRWLK 5377

QY 129 ----- 128
Db 5378 GTQEITGDDRFELIKDGTKHSMVIKSAAPFEDEAKYMFEAEDKHTSGKLIIGIRLKLFTLP 5437
QY 129 ----- 128
Db 5438 LKDVTAKEKESAVFTVELSHDNIRVKFNQDORLHTTRSVSMQDEGKTHSITFKDLSIDD 5497
QY 129 ----- 128
Db 5498 TSQIRVEAMGMSSEAKLTVLEGDPIFTGKLQDYTGVEKDEVILQCEISKADAPVKWFKDG 5557
QY 129 ----- 128
Db 5558 KEIKPSKNAVAKDGKRRMLILKALKSDIGQYTCDCGTDKTSGLDIEDREIKLVRPLH 5617
QY 129 ----- 128
Db 5618 SVEVMEETARPETEISEDDIHANNKLKGEALLQTPDCEIKEEGKIHSLSVLHNCRLDQTG 5677
QY 129 ----- 128
Db 5678 GVDFOAANVYKSAHLRVKPRVIGLLRPLKDVTVTAGETATFDCELSYEDIPVEWYLGKK 5737
QY 129 ----- 128
Db 5738 LEPSDKVVRSEGVHTLTLRDVKLEDAGEVOLTAKDFKTHANLFEVKEPPVEFTKPLEDQ 5797
QY 129 ----- 128
Db 5798 TVEEGATAVLECEVSRENKAVKFKNGTEILSKKYEIVADGRVRKLVITHDCTPEDIKTY 5857
QY 129 ----- 128
Db 5858 TCDAKDFTSCNLNVVPPHVEFLRPLTDLVREKEMAREFCELSRENKVKWFKDGAIEK 5917
QY 129 ----- 128
Db 5918 KGKDYIISKGAVRILVINKCLLDDAEYSCVETARTSGMLTVLEEEAVFTKNLANIEV 5977
QY 129 ----- 128
Db 5978 SETDTIKLVCEVSKPGEVIVYKGDDEIETGRYEILTEGRKRILVIONAHLEDAGNYC 6037
QY 129 ----- 128
Db 6038 RLPSSRTDGKVKVHELAAEFISKPONLEILEGKAEFVCSISKESFPQMKRDDKTLESG 6097
QY 129 ----- 128
Db 6098 DKYDVIADGKKRVLVVKDATLQDMGTYYVMVGAARAAHLTVIEKLRIVVPLKDTVRKEQ 6157
QY 129 -----VYTL----- 132
Db 6158 QEVFVCEVNTGAKAKWFRNEEAFDSSKYLILQKDLVYTLRIRDAHLDOQOANYSLT 6217
QY 133 ----- 132
Db 6218 NHRGENVSAANLIVEEBDLRIVEPLKDIETMEKKSVTFWCKVKNRLNVTLKWTKNGEVP 6277
QY 133 ----- 132
Db 6278 FDNRSYRVDKYKHLTIKDCGFPDEGEYIVTAGODKSVAELLIIIEAPTEFVEHLEDQTV 6337
QY 133 ----- 132
Db 6338 TEFDDAVFSCQLSREKANVKWYRNGREIKGKFKERDGSIHRLI IKDCLDDECEYAC 6397
QY 133 ----- 132
Db 6398 GVEDRKSARLFEVEIPVEIIRPPQDILEAPGADVVFVLAELNKKDKVEVQWLRNNMVVQG 6457

QY	133	-----	132
Db	6458	DKHQMSEGIHRLQICDIKPROGEYRFTAKDEARAKLELAAAPKIKTADQDLVVDVG	6517
QY	133	-----	132
Db	6518	KPLTMVVPYDAYPKAAENFENEPLSTKTIDTTAEQTSFRILEAKKGDKRYKIVLQNK	6577
QY	133	-----	132
Db	6578	HGAEGFINLKVINDVPGVVRNLEVTFTDGEVSLAMEEPLTDGGSKIIGYVVERDIKRR	6637
QY	133	-----	132
Db	6638	TWVLATDRAESCEFTVTGLQGGVEYLFVRSARNRVTGEPVETDNPVEARSKYDVPGP	6697
QY	133	-----	132
Db	6698	LNVITIDVNRFGVSLTWPPEYDGGAEITNYVIELRDKTSIRWDTAMTVRAEDLSATVTD	6757
QY	133	-----	132
Db	6758	VVEGQEYFRVRAQNRIGVCKPSAATPFVKVADPIERPSPVNLTSDDQTQSSVOLKWE	6817
QY	133	-----	132
Db	6818	PLKGGSPILGYIIERCEECKDNWIRCNMKLVPELTYKVTGLEKGNKLYRVAENKAGV	6877
QY	133	-----	132
Db	6878	SDPSEILGPLTADDAFVEPTMDLSAFKDGLEIVPNPITILVPSTGYPRPTATWCFDGV	6937
QY	133	-----	132
Db	6938	LETGDRVKMKTLSAYAEVLSPERSDKGIYTLKLENRVKTSIGEDVNVIAEPSAPKEL	6997
QY	133	-----	132
Db	6998	KFGDITKDSVHLTWEPPDDGGSLTGYVVEKREVSRTKTKYMDFTDLEFTVPLVQG	7057
QY	133	-----	132
Db	7058	KEYLFKVCARNKCGPEPAYVDEPNMSTPATVDPDENVKWRDRTANSIFLTWDPKND	7117
QY	133	-----	132
Db	7118	GGSRIGYIVERCPGSDKHWACGEPVAETKMEVTGLEGKWYAYRVKALNRQASKPSR	7177
QY	133	-----	132
Db	7178	PTEEIOAVDTQEAPEIFLDVKLLAGLTVKAGTKIELPATVTGKPEPKITWTKADMILKQD	7237
QY	133	-----	132
Db	7238	KRITIENVPKSTVTIVDSKRSRTGYII EAVNVCGRATAVEVNVLDKPGPAAFDITD	7297
QY	133	-----	132
Db	7298	VTNESCLLTWNPPRDDGSKITNYVVERRATDSEVHKLSSYTKDTNFKATKLIPNKEYI	7357
QY	133	-----	147
Db	7358	FRVAENMYGVEPVQASPIKAYQFDPGPPTRLEPSDITKDAVLTWCEPDDGGSPI	7417
QY	148	-----	147
Db	7418	TGYVVERLDPTDKWRCNMKPVKDTTYRVKGLTNKKYFRVLAENLAGPGKPSKSTP	7477
QY	148	-----	147
Db	7478	ILIKDIPDPWPGKPTVKDVGKTSVRLNWKPEHGGAKIESYVIEMLKTGTDEWVRVA	7537
QY	148	-----	147

Db	7538	EGVPTTQHLPLGLMEGOEYSFRVRAVNKAGESESESDPVLCREKLYPPSPRWLEVIN	7597
QY	148	-----	147
Db	7598	ITKNTADLKWTVPKDGSGPITNYIVEKRDVRRKGWQTVDTTVKDKTKCTPTPLTEGSLYV	7657
QY	148	-----	147
Db	7658	FRVAENAIQSDYTEIETEDSVLAKDTFTTGPYPYALAVDVTKRHVDLKWEPKNDGGRP	7717
QY	148	-----	147
Db	7718	IQRVIEKKERLGRWVKAGKTAGDCNFRVTDVIECTEVOFQVRAENAGVHPSBTE	7777
QY	148	-----	147
Db	7778	ILSIEDTSPSPPLDLHVTDAGRKHIAIAWKPEKNGGSPIIIGYHVEMCPVGTEKMRV	7837
QY	148	-----	147
Db	7838	NSRPDKLKFVEGVVPDKEYVLRVAVNAIGVSESEISENVVAKDPDCKPTIDLETH	7897
QY	148	-----	147
Db	7898	DIIVIEGEKLSIPVPFRAPVPPTVSWHKDGKEVKASDRLTMKNDHISAHLEVPKSVRADA	7957
QY	148	-----	147
Db	7958	GIYTITLENKLSATASINVKVIGLPGCPCKDIKASDITKSKCLTWEPPEFDGTPILHY	8017
QY	148	-----	147
Db	8018	VLERREAGRTRYIPVMSGENKLSWTVKDLIPNGEYFRVAVNAVKVGGEYIELKNPVAQ	8077
QY	148	-----	147
Db	8078	DPKQPPDPVDVEVHNPTAEAMTITWKPLYDGGSKIMGYIIEKIAKEERKRCNEHLV	8137
QY	148	-----	147
Db	8138	PILTYTAKGLEEGKEYQFVRAENAAIGSEPSRATPPTKAVDPIDAPKVILRTSLEVKRG	8197
QY	148	-----	147
Db	8198	DEIATDASISGSPYPTITWIKDENVIYVEEIKKRAAPLVRRRKGEVQEEPFVLPQLRL	8257
QY	148	-----	147
Db	8258	SIDNSKKGESQLRVRDSLDPDHGLYMIKVENDHGIAKAPCTSVSLDTPGPPINFVEDIR	8317
QY	148	-----	147
Db	8318	KTSVLCKWEPPLDDGSEIINYTEKKDKTKPQSEWIVVTSTLRHCKYSVTKLJEGEYL	8377
QY	148	-----	147
Db	8378	FRVAENRFGPGPCVSKPLVAKDPFGPDAPDKPIVEDYTSNSMLYKNWNEPKDNGSPIL	8437
QY	148	-----	147
Db	8438	GYWLEKREVNSTHWSRVNKSILLNALKANVDLLEGLTYFVRVCAENAAAGPKFSPSPDK	8497
QY	148	-----	147
Db	8498	TAHDPISPGPPIPRVTDTSSTTIELEWEPFANGGGEIVGYFVDKQLVGTNWSRCTEK	8557
QY	148	-----	147
Db	8558	MIKQYTVKEIREGADYKLRVSAVNAAGEGPGTQPTVVAEPQEPAPVELDVSVKGGI	8617
QY	148	-----	147

Db 8618 QIMAGKTLIPAVTGRPVPTKVWTKKEGELDKDRVIDNVGTSKSELIIKDALKDHRG 8677
QY 148 ----- 147
Db 8678 VITATNSCGSKFAAARVEFDVPGVLDLKPVTNRKMKLLNWSDEDDGGSEITGFIIE 8737
QY 148 ----- 147
Db 8738 RKDAKHTWROPIETERSKCDITGLLEGOEYKFRVIAKNKFCGPPVEIGPILAVDPLGP 8797
QY 148 ----- 147
Db 8798 PTPERLTYTERTKSTITILDMKEPRSGNGSPIQGYIIIEKRHRDKPDFERVNKRCLPTTSF 8857
QY 148 ----- 147
Db 8858 LVENLDEHOMYEFVRKAVNVEISEPSPLNVVIOQDEVPTIKRLSVRGDTIKVKAGE 8917
QY 148 ----- 147
Db 8918 PVHIPADVTGLPMPKIENKNETVIEKPTDALQITKEEVSRSSEAKTELSIPKAVREDKGT 8977
QY 148 ----- 147
Db 8978 YTVTASNRLGSVFRNVHVEYDRPSPPRNLAVTDIKAESCYLTWDAPLDNGGSEITHYVI 9037
QY 148 ----- 147
Db 9038 DKRDASRKAHEEVTNTAVERKYGIIWKLI PNGQYEFVRVAVNKYGISDECKSKVVIQD 9097
QY 148 ----- 147
Db 9098 PYRLPCPGKPKVLARTKGSMLVSWTPTPLDNGSPITGVWLEKREEGSPYNSRVSRAPIT 9157
QY 148 ----- 147
Db 9158 KVLKGVFNVRLLGVKYQFRAMA INAAIGIGPPSEPDPVAGDPIFPFGPPSPCEVK 9217
QY 148 ----- 147
Db 9218 DKTKSSISLGNKPPAKDGGSPKIGYIVEMQEBGTTDKRVNPDKLIITCECVVFNKEL 9277
QY 148 -----CLV----- 150
Db 9278 RKYFRVKA VNAEGESEPDTTGEIPATDIOEPEVFIDIGAQCCLVCKAGSOIRIPAVI 9337
QY 151 ----- 150
Db 9338 KGRPTPKSSWEPDGAKKAMKDGVDHIDIPEDAQLETAENS SVIIIPCKRSHTGKYSITAK 9397
QY 151 -----KGF----- 153
Db 9398 NKAGQKTANCRVKVMDVPGPKDLKVS DITRGCSRLSNKMPDDDDGDRIGYVIEKRTID 9457
QY 154 -----YPSD- 157
Db 9458 GXAWTKANPCDGTTFVVPDLLEQQYEFVRVAENRFGIPGVETIQRTTARDPIYPPDP 9517
QY 158 ----- 157
Db 9518 PIKILGLITKNTVHLSWKPKPKNDGGSPVTHYIVECLAWDPTGTTKEAWRQCNKRDEEL 9577
QY 158 ----- 157
Db 9578 QFTVEDLVEGGEYEFVRKAVNAAGSKPSATVGPCDQCPDMPSPSIDLKEFMEVEEGTNV 9637
QY 158 ----- 157
Db 9638 NIVAKIGVPEPTLTFWFKAPKPKONKEPVLVDTHVNLVDDTCTLVIPQSRSDTGLY 9697
QY 158 ----- 157
Db 9698 TITAVNNLGTASKEMRLNLVLRGPPVGPINKFESVSADQMTLSWPPPKDGGSKITNVYI 9757

QY 158 ----- 157
Db 9758 EKREANRKTWVHVSSEPKECTYTI PKLLEGHEYVPRIMAQNKYGICEPLDSEPETARNLF 9817
QY 158 -----IAVEWE----- 163
Db 9818 SVPGAPDKPTVSSVTRNSMTVNWEPEYDGGSPVTGYWLEMKDTSKRWRVNRDPIKAM 9877
QY 164 ----- 163
Db 9878 TLGVSYKVTGLIEGSDYQFRVYAINAAGVGPASLPSPATARDPIAPGPPFPKVTDMTK 9937
QY 164 ----- 163
Db 9938 SSADLEWSPLKDGSKVTGYIYEYKEEGKEWEGKDKVEVRGTLVVTGLKEGAFYKFR 9997
QY 164 ----- 163
Db 9998 VSAVNIAGIGEPGEVTDVIEMKDRLVSPDLQLDASVRDRIVVHAGGVIRIIAIVVSGKPPP 10057
QY 164 ----- 163
Db 10058 TVTWMNERTLPQEATIIETAISSMVINKCORS HQGVYSLLAKEAGERKKTIIIVDVL 10117
QY 164 -----SNGOPENNY----- 172
Db 10118 VPGPVGTPLAHLNLTNESC KLTWFSPEDDGSGSPITNYVIEKRESORRAWTPVYTYVTRON 10177
QY 173 ----- 172
Db 10178 ATVQGLIQGKAYFRIAAENSIGMPFVETSEALVIREPITVPERPEDLEVKEVTNTVT 10237
QY 173 ----- 172
Db 10238 LTWNP KYDGGSEIINYLESLIGTEKPHKVTNDNLLSRKYTVVGLKEGDTYEVRSVAV 10297
QY 173 ----- 172
Db 10298 NIVGOGKPSFCTK PITCKDELAPPTLHLDFRDKLTIRVGEAFALTGRYSGKPKVSNPK 10357
QY 173 -----KTP----- 176
Db 10358 DEADVLEDDRTHIKTTPATLAL EKIKA RSDSGKYCVVVENSTGSRKGFCQVNVYDRGP 10417
QY 177 ----- 176
Db 10418 PVGPVSFDEVT KDYMVISWKPPPLDGGSKITNYIIIEKKEVGKDWMPVTSASAKTCKVS 10477
QY 177 ----- 176
Db 10478 KLEGGDYIFRIHAENLYGISDPLVSDSMKAKDRFPVDPADPQPIVTEVTKDSALVTWNK 10537
QY 177 ----- 176
Db 10538 PHDGGKPIITNYILEKRETKSRWARVTKDPIHPYTKFRVPDLLEGCQYEFVRSAE NEIGI 10597
QY 177 ----- 176
Db 10598 GDPSPKPVFAKDIAPKPPVPNPEADITTCNSVDLTWQPPRHGGSKILGYIYEQKV 10657
QY 177 ----- 176
Db 10658 GDEWRANRHTPESC PETKYKVTGLRDGQTYKFRVLAVNAAGESDPAHVPEPVLVKDRLE 10717
QY 177 ----- 176
Db 10718 PPELILDANMAREQHIKVGD TLRLSAIIKGVFPKVTWKEDRDAPTARIDVTPVSGKL 10777
QY 177 ----- 176
Db 10778 EIRNAAHEDGGIYSLTVENPAGSKTVSVKVLVLDKFGPPDRDLEVSEIRKDCYLTWKEPL 10837

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:21 ; Search time 28.9412 Seconds

(without alignments)
1132.630 Million cell updates/sec

Title: 2LINK7

Perfect score: 1334

Sequence: 1 MDKTHPCPCPAPPELLGSPS.....GKDWLKAFYDKVAEKLEAF 246

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq.-101002:*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1246	93.4	247	21	AA16958
2	1246	93.4	247	23	AB173411
3	1246	93.4	248	21	AB17953
4	1246	93.4	248	23	AB173421
5	1246	93.4	268	21	AA16959
6	1246	93.4	268	23	AB173412
7	1246	93.4	269	21	AA196531
8	1246	93.4	731	22	AA152156
9	1246	93.4	741	22	AA152159
10	1245	93.3	374	19	AA183963

11	1245	93.3	374	19	AA19075
12	1245	93.3	401	22	AA172922
13	1245	93.3	401	22	AA180904
14	1245	93.3	423	21	AA186633
15	1245	93.3	441	21	AA186692
16	1245	93.3	689	21	AA196780
17	1245	93.3	729	22	AA152158
18	1245	93.3	739	22	AA152161
19	1244	93.3	652	19	AA18650
20	1243	93.2	633	21	AA184965
21	1242	93.1	252	21	AA17955
22	1242	93.1	252	23	AA173423
23	1242	93.1	689	21	AA196779
24	1241	93.0	426	21	AA186694
25	1241	93.0	448	21	AA186694
26	1240	93.0	379	19	AA183962
27	1240	93.0	379	19	AA19073
28	1240	93.0	379	19	AA19074
29	1240	93.0	651	18	AA126649
30	1240	93.0	692	18	AA126650
31	1239	92.9	248	21	AA17951
32	1239	92.9	248	23	AA173419
33	1239	92.9	253	21	AA16964
34	1239	92.9	253	23	AA173415
35	1239	92.9	277	21	AA16967
36	1239	92.9	277	23	AA173418
37	1239	92.9	381	22	AA181972
38	1239	92.9	582	22	AA181987
39	1239	92.9	582	22	AA181991
40	1239	92.9	583	22	AA183156
41	1238	92.8	228	21	AA16955
42	1238	92.8	228	21	AA196529
43	1238	92.8	228	22	AA196933
44	1238	92.8	228	23	AA181074
45	1238	92.8	228	23	AA173410

ALIGNMENTS

RESULT 1	
AA16958	
ID	AA16958 standard; Protein; 247 AA.
XX	
AC	AA16958;
XX	
DT	31-OCT-2000 (first entry)
XX	
DE	Fc-TMP protein sequence SEQ ID NO:6.
XX	
KW	Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW	autoimmune disease; cytostatic; antitumorigenic; thrombolytic; VEGF;
KW	immunopressure; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW	inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW	cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW	vascular endothelial growth factor; matrix metalloproteinase;
KW	asthma; thrombosis; pharmaceutical.
XX	
OS	Homo sapiens.
OS	Synthetic.
PN	WO200024782-A2.
XX	
PD	04-MAY-2000.
XX	
PF	25-OCT-1999; 99WO-US25044.
XX	
PR	23-OCT-1998; 98US-0105371.
XX	
PA	22-OCT-1999; 99US-0428082.
XX	
PI	(AMGE-) AMGEN INC.
	Felge U, Liu C, Cheetham J, Boone TC;

Recombinant human
Human met-Fc (lack
Human metFc-
Fc-huA2P-1 (114-28
Fc-huA2P-1 (95-281
Ang-2-PD-Fc-Fc fus
Humanised HMG-1 h
Humanised HMG-1 h
Heavy chain of hma
Amino acid sequenc
Fc-VEGF antagonist
Fc-VEGF antagonist
Ang-1-PD-Fc-Fc fus
Fc-huA2P-1 (120-29
Fc-huA2P-1 (99-291
Recombinant human
Recombinant human
Recombinant human
Chimeric receptor
Fc-TNF-alpha inhib
Fc-TNF-alpha inhib
Fc-EMP protein seq
Fc-EMP mimetic pep
Fc-EMP-EMP protein
Fc-EMP-EMP nucleic
Ganglioside GD2 sp
Ganglioside GD3 sp
Ganglioside GM2 an
Human IgG1 Fc prot
Human IgG1 Fc chai
Human IgG1 Fc regi
Human IgG1 Fc. Ho
Human immunoglobul

Db 24278 SSAISVTTTCRIVNLNENAGYYFRVSAQNTFGISDPLEVSSVVIKSPFEKPGAGPKPTIT 24337
QY 205 -----SCSV----- 208
|||
Db 24338 AVTKDCVVAWKPPASDGGAKIRNYLEKREKKONKWSITTEIRETVFSVKNLIEGLE 24397
QY 209 ----- 208
Db 24398 YEFVRKCNLGESEWSEISEPITPKSDVPDIOAPHFKEELNRLNVRYSQSNATLVCKVTGH 24457
QY 209 ----- 208
Db 24458 PRPIVKWYRQGEIADGLKYRIOEPKGYHQLIITASVTDATVYOVVRATNOGGSVST 24517
QY 209 ----- 208
Db 24518 ASLEVEVPKAIHLPKTLEGMAVHALRGEVWSIKIPFSGKPDVPVITWQKGQDLIDNNGHY 24577
QY 209 ----- 208
Db 24578 QVIVTRSFVSLVPPNGVERKDAGFYVYCAKNRFGIDQKTVELDVADVPDPGRGVYSDAS 24637
QY 209 ----- 208
Db 24638 RDSVNLWTPEASDGGSKITNIVEKCATTAERWLRVGQARETRYVINLFKTSYQFRV 24697
QY 209 ----- 208
Db 24698 IAENKFGLSKPSPEPTITKEDKTRAMNYDEEVDREVSMTKASHSTKELYKYMIA 24757
QY 209 -----MHEA- 212
|||
Db 24758 EDLGRGEGVHRCVETSSKKTMAKFKVKGTDQVLVKKESILNIAHRNHLHESF 24817
QY 213 -----LHNH----- 216
|||
Db 24818 ESMEELVLFETISGLDIFERINTSAFELNREIVSYHVQVCEALQFLSHNIGHDIRP 24877
QY 217 ----- 216
Db 24878 ENIIYOTRRSSTIKIIEFGQARQLPGDNFRLLFTAPEYAYEVHQHDVVSTATDMWSLG 24937
QY 217 -----YT----- 218
|||
Db 24938 TLVVLLSGINPLAETNOQIIENINNAEYTFDEAFKEISIEAMDFVDRLLVKKRKSAM 24997
QY 219 -----OXS 221
|||
Db 24998 TASEALQHPWLKQIERSVSTKIVIRTLKHRRYYHTLIKKDLNMVVSARISCGAITSQKG 25057
QY 222 LSL----- 225
|||
Db 25058 VSAKVVKVASIEIGPVSGOIMHABEGGHVKYVCKIENYDQSTQVWTWFGVRQLENSEK 25117
QY 226 ----- 225
Db 25118 YEITYEDGVAILLYVKDITKLDDGYRCKVYVNDYGEDSSYAEFLVKGREVDYCYRRUMK 25177
QY 226 ----- 225
Db 25178 KIKRRDTMRLERPPEFTLPLYNKTAUYGVNVRGVTITVHPEPHVTWYKSGQIKPGD 25237
QY 226 ----- 225
Db 25238 NDKKYTFESDKGLYQLTINSVTDDDAEYTVVARNKYGEDSCKAKLTVLHPPTDSTLR 25297
QY 226 ----- 225
Db 25298 PMFKRLLANAQEGQSGVCFEIRVSGIPPTLKWEDGQPLSIGPNIEIIEHGLDYALH 25357
QY 226 ----- 225
Db 25358 IRDTLPEDTGYRVTAINTAGSTSCQAHLOVERLYKKQEFKSEHEHRHVQKIDKTUR 25417

QY 226 ----- 225
Db 25418 MAEILSGTESVPLTQVAKREALREAAVLYKPAVSTKTVKGEFRLEIEEKKEERKLMPYDV 25477
QY 226 -----PGK----- 228
|||
Db 25478 PEPRKYKOTTIEEDORIKOFVPMDSMKWKIKINDQYEMFGKLDKRVVQVKRKRIRLSWBQ 25537
QY 229 ----- 228
Db 25538 FYVMPPLPRITDQYRKPWRIPKLSODDLEIVRPARRTPSPDYDFYRPRRSLSGDISDEE 25597
QY 229 ----- 228
Db 25598 LLLPIDDYLAMKRTEERLURLEELGFGSPPSPHFFELSLURYSPPQAHVKEET 25657
QY 229 ----- 228
Db 25658 RKNFRYSTYHIPTKAEASTSYAELRERHAQAAAYRQPKQRORIMAREDEBELLRPVTTQH 25717
QY 229 ----- 228
Db 25718 LSEYKSELDPMSEKSRKSRQRREVTTEIEEBEYFISKHAQRESSSASRLLRRRS 25777
QY 229 ----- 228
Db 25778 LSPTYIELMRPVSELIRSRQPAEEYEDDTERRSPTTPRSPSPVSSERSLSRFRS 25837
QY 229 ----- 228
Db 25838 ARFDIFSRYSKMAALKTKQTSERKEYVLSQQPFTLDHAPRITLRMRSHRVPQGNTRFI 25897
QY 229 ----- 228
Db 25898 LNVQSKPTAEVKWYHNGVELQESSKIHNTSGVLTLEILDCHTDDSGTYRAVCTNYKE 25957
QY 229 ----- 228
Db 25958 ASDYATLDVTGGDYTYASQRRDEEVPSPFELTRTEAYAVSPFKKTSMEASSSVREV 26017
QY 229 ----- 228
Db 26018 KSQMTETRESLSYEHSAEMKSAALEKESLEKSTTRKIKTTLAARILTKPRSMTVYE 26077
QY 229 -----DWL----- 231
|||
Db 26078 GESARFSCDTEGPEPVTWLRKGOVLSTSAHQVTTTKYSTFEISSVQASDEGNYSV 26137
QY 232 ----- 231
Db 26138 VENSEGQBAEFTLTIQKARVTEKAVTSPPRVKSPPEPRVKSPKAVKSPKRVKSPSPHPK 26197
QY 232 ----- 231
Db 26198 AVSPTETKTPREKVOHLPVSAPPKITQFLKABASKEIAKLTCVWESSVLRAKEVTWYKD 26257
QY 232 -----KAF 234
|||
Db 26258 GKLLKENGHFQHYSDAGTYELKINNLTESDQGEYVCEISGEGTSKTNLQPMGOAFKSI 26317
QY 235 YDKVAE----- 240
|||
Db 26318 HEKVSISSETKKSDQKTTESTVTRKTEKPAPEPISSKPVIVTGLQDQTTVSSDSVAKFAVK 26377
QY 241 ----- 240
Db 26378 ATGEPRPTAINTKDGKAITQGGKYLSEDKGGFFLEIHKTDTSGLYTCYTKVNSAGSVS 26437
QY 241 -----KLKEA 245
|||
Db 26438 SSCKLTIKAIDTEAKQKVSTQKTSETPOKAVVQBEISQKALRSEIKMSEA 26490

Qy 183 ----- 182
Db 22118 EFRVAKNDAGYSEPREAFSSVIIKEPIEPTADLTGITNLIITCKAGSPFTIDVPISGR 22177
Qy 183 -----GSFFL----- 187
Db 22178 PAKVTWLEEMRLUKETDRVSITTTKORTTLTVKDSMRGDSGRYFLTILENTAGVKTFEVT 22237
Qy 188 ----- 187
Db 22238 VVIGRPGVPTGPIEVSSVSAESCVLSWGEPKDGGEITNVIYVEKRESGTTAMOLVNS 22297
Qy 188 ----- 187
Db 22298 VKRTQIKVTHLTMYEYFRVSSENREFGVSKPLESAPIIAHPPVPPSAPTRPEYHVSA 22357
Qy 188 ----- 187
Db 22358 NAMSIRWEEPYHDGSKIIGYWEKKEKERTILWVKENKVPCECNKYVTGLVEGLEXYQR 22417
Qy 188 ----- 187
Db 22418 TYALNAGVSKASEASRPIMAQNPVDAPGRPEVTDTRSTVSLIWSAPAYDGGSKVGYI 22477
Qy 188 ----- 187
Db 22478 IERKPVSEVGDRWLKCNVTIVSDNFTVTALSEGDTYEFRLAKNAGVISKSESTGP 22537
Qy 188 ----- 187
Db 22538 VTCREYAPPAELDARLHGLDVTIRAGSDIVLDAAVGGKPEPKIITWTKGDKELDCEK 22597
Qy 188 ----- 187
Db 22598 SLOVTGKRATAVINPCDRSDSGKYTLTVKNASGTVKAVSMVKVLDSPGCKLTVSRVTO 22657
Qy 188 ----- 187
Db 22658 EKCTIANSLOEDGGAETHYIVERRETSRLNWIVVEGECTLSYVVTRLIKNEYIFRV 22717
Qy 188 ----- 187
Db 22718 RAVNKYGPVSEPIVARNSTIPSPGPIPEVGTGKEHIIQWTKPESDGGNEISNY 22777
Qy 188 ----- 187
Db 22778 LVDRREKESLRWTRVNKDYYVVDTRLKVTSLMEGCDYQFRVTAVNAAGNSEPSESNFIS 22837
Qy 188 ----- 187
Db 22838 CREPSYTPGPSAPRVVDTTKHSISLAWTKPMYDGGTDIVGYVLEMQEKOTDQWYRVHTN 22897
Qy 188 ----- 187
Db 22898 ATIRTEFTVPLKMGOKYSPRAAVNVKGMSEYSESAIEPVERIEIPDELEADLKK 22957
Qy 188 -----YSKLTVDK----- 195
Db 22958 TVTTRAGASLRMLVMSGRPVPTWTSKQIDLASRAIIDTTESYSLLIVDKVNRDAGK 23017
Qy 196 ----- 195
Db 23018 YTIEAENQSGKKSATVLVKVYDTPGPCPSVKVKEVSRDSVTITWIEPTIDGGAPINNYI 23077
Qy 196 ----- 195
Db 23078 EKREAAAMFAKTVTTCCKSTLYRISGLVEGTMHYFRVLPENIYGIGPEPCTSDAVALYSEV 23137
Qy 196 ----- 195
Db 23138 PLVPAKLEVVDVTKSTVTLAWEKPLXDGGSLRTGYVLEACKAGTERMKNVVTLPKPTVLEH 23197
Qy 196 ----- 195

Db 23198 TVTSLNEGEQYLFRIQAQNEKGVSEPRETVTAFTVQDLRVLPTIDLSTMPKQTIHVPAGR 23257
Qy 196 ----- 195
Db 23258 PVELVIPIAGRPPPAASWFFAGSKLRESERVTVETHTKVAKLTIRETTRDTGTYTLELK 23317
Qy 196 ----- 195
Db 23318 NVTGTTSETIKVIIIDKPGPPTGPIKIDEIDATISITISWEPPELGGAPLSGYVVEORDA 23377
Qy 196 ----- 195
Db 23378 HRCGLPVSSEVSTRSTFKFTRLTEGNEVYFRVAATNREFIGSYLQSEVIECRSSIRIPGP 23437
Qy 196 ----- 195
Db 23438 PETLIQIFDVSRDGMTLTWYPPEDDQGSQVTGYIVERKEVRADRWVRVKNKVPVMTYRST 23497
Qy 196 ----- 195
Db 23498 GLTEGLEYHRVTVAINARGSGKPSRPSKPIVAMPDPIAPPKQPONPRVTDTRTSVSLAWS 23557
Qy 196 ----- 195
Db 23558 VPDEGGSKVTGYLIEMQKVDOHEWTKCNTPTTKIREVTLTLHPQGAERYRVLACNAGG 23617
Qy 196 -----SRWQOG----- 201
Db 23618 PGPAEYPGTVKVTEMLEYPDYELDERYQEGIFVRQGGVIRLTIPIKGKPPICKWTKEG 23677
Qy 202 ----- 201
Db 23678 QDISKRAMIATSETHTELVIKEADRGDSGYDVLNKKCGKAVYIKVRVIGSPNSPEGP 23737
Qy 202 ----- 201
Db 23738 LEYDDIOQVRVRSWRPPADDDGADILGYILERREVKAANYTIDSRVGTSLVVKGLKE 23797
Qy 202 ----- 201
Db 23798 NVEYHFRVSAENORFISGPLKSEPVTPKPLNPPEPPSNPPEVLDVTKSSVLSWSRPK 23857
Qy 202 ----- 201
Db 23858 DGGSRVTGYIERKETSTOKVVRHNKTQITTTWYTVGLVPDAEYQFRIIAQNVDGLSE 23917
Qy 202 ----- 201
Db 23918 TSPASEPVCKDPDPKPSQPGLELILSISKDSVTLOWEKPECDDGKELGCVWVEYRQSGD 23977
Qy 202 ----- 201
Db 23978 SAWKSNKERIKDKQFTIGLLEATEYFRVFAENETGLSRPRRTAMSIKTLTSGEAPG 24037
Qy 202 ----- 201
Db 24038 IRKEMKDVTTKLGEEAOLSCQIVGRPLPDIKWYRFGKELIOSRKYKMSDGRHTLTVMT 24097
Qy 202 ----- 201
Db 24098 EEQDEGEVYTCIATNEVEGETSSKLLQLQATPQFHPGCPYPLKEYYGAVGSTLRHLHVYIG 24157
Qy 202 -----NVF----- 204
Db 24158 RVPAMTWFHQKLLQNSENITIENTEHTYHLVMKNVQRKTHAGKYKQLSNVEGTDAI 24217
Qy 205 ----- 204
Db 24218 LDVEIODPKPTGPIVIEALLKNASVSWKPPADGGSGMITNVVVEKBAKEAGQOLV 24277
Qy 205 ----- 204

Db 19898 WLKPEHGGSRIVHYVVEALEKQGKNWKCVAKSTHHVVSGLRENSEYFFRVFAENOAG 19957
QY 176 ----- 175
Db 19958 LSDPRELLPLVLIKEOLEPEIDMKNFSPSHYVVRAGSNLKYVDIPISGKPLPKVTLSDRG 20017
QY 176 ----- 175
Db 20018 VPLKATMRENTETAENLTINLKESVTADAGRYEITAANSOGTTKAFINIVVLDPRGPPT 20077
QY 176 ----- 175
Db 20078 GPWISDITEESVTLKNEPPKYDGGSQVTNYILLKRETSTAVWTEVSATVARTMKVMKL 20137
QY 176 ----- 175
Db 20138 TTGEYQFRIKAENRFGISDHIDSACVTVKLPYTTTPGPSTPWVTVNVTRESITVGHEPV 20197
QY 176 ----- 175
Db 20198 SNGSAVGYHLEMKDRNSILWQKANKLVIRTHFKVTTISAGLIYEFVRYAENAGVGK 20257
QY 176 ----- 175
Db 20258 PSHPEPVLADACEPPRNVRTIDISKNSVLSWOQAPADGGSKITGYIVERDLDGGRW 20317
QY 176 ----- 175
Db 20318 TRASFTNVTTQTTISGLTONSQYEFVFNARNAVGSINPSEVVGPIITCIDSYGGPVIDL 20377
QY 176 ----- 175
Db 20378 PLEYTEVKYRAGTSVKLRAGISGKPAPTIEWYKDDKELQTNALVCVENTTDLASILND 20437
QY 176 ----- 175
Db 20438 ADRLNSGCYELKLRNMAASATIRVQILDKPGPGPIEFKVTVAEKTILLWRPPADDG 20497
QY 176 ----- 175
Db 20498 GAKITHYIIVEKRETSRVVMSVSEHLEECIITTKIIGNEYIFRVAVNKYIGIGEPLES 20557
QY 176 ----- 175
Db 20558 DSVAKNAFVTPGPGIPEVTKITNSMTVWVSRIADGGSDISGYFLEKRDKKS LGWFK 20617
QY 176 ----- 175
Db 20618 VLKETIRDTRQKVTGLTENSQYRVCVAVNAAGQGFSEPFSEFYKAAIDPDPGPPAKIR 20677
QY 176 ----- 175
Db 20678 IADSTKSSITLGWSRPVYDGSASVTGYVVEIRQGEEETVSTRKEVTRTEYVVSNLKP 20737
QY 176 ----- 175
Db 20738 GYNYFRVSAVNCAGGEPPIEMNEPVQAKDILEAPEIDLVALRTSVIAKAGEDVQVLIP 20797
QY 176 ----- 175
Db 20798 FKGRPPPTVTRKDEKLGSDARYSIENTDSSLLTIPQVTRNDTGKYLITTIENGVEPK 20857
QY 176 ----- 175
Db 20858 SSTSVKVLDTPAACQKLOVKHVSRGVTLLWDPPLIDGSPINIVYIEKRDATKRTWSV 20917
QY 176 ----- 175
Db 20918 VSHKCSSTFKLIDLSEKTPPFVFRVLAENEIGIGECETTEPVKAAEVPAPTRDLSMKDS 20977
QY 176 ----- 175
Db 20978 TKTSVILSWTKPFDGGSVITEYVVVERKKGQBTWSHAGISKTCIEIVSOLKEQSVLEER 21037

QY 176 ----- 175
Db 21038 VFAKNEKGLSDPVTIGTITVKELIITPEVDLSDIPGAQVTVRIGHNVHLELPYKGPKPS 21097
QY 176 ----- 175
Db 21098 ISWLKDGFLPKESEFVRFSTENKITLSIKNAKHEGGKYTVILDNAVCRIAVPITVTL 21157
QY 176 ----- 175
Db 21158 GPPSKPKGPIREDEIKADSVILSWDPEDNGGGEITCYISIEKRETSQTNWKMVCSSVART 21217
QY 176 ----- 175
Db 21218 TEKVPNLKDAEYQFVRAENRYGSQPLVSSIIVAKHQFRIPGPGKPVINYNTSDGMS 21277
QY 176 ----- 175
Db 21278 LFWDAVPYDGGSEVTGFHVEKKERNLSILWQKNTSPISGREYRATGLVEGLDYQFRVYAE 21337
QY 176 ----- 175
Db 21338 NSAGLSSPDSKFTLAVSPVDPGTPDYIDVTRITILKNPPLRDGGSKIYGYSIER 21397
QY 176 ----- 175
Db 21398 QGNERWVRNCFDVSQCQVTVTGLSPGDRYEFRIIARNAVGTISPSPSSQSGIIMTRDENV 21457
QY 176 PPV----- 178
||:
Db 21458 PPVIFGPEYFDGLIJKSGESURIKALVOGRVPVPRVTFKDGVEIEKRMNMEITNVLGST 21517
QY 179 ----- 178
Db 21518 SLFVRDATRDRGVTVVEAKNAGSASAKELVKVQDTPGVVGPFRFTNITGEKMTLWWD 21577
QY 179 ----- 178
Db 21578 APLNDGCAPITHYIIEKRETSRLAWALIEDKCEAQSYTAIKLINGNEYQFRVSANVKFGV 21637
QY 179 ---LDS----- 182
||||
Db 21638 GRPLSDPVPVAIQVTVDPAGIPEPSNITGNSIILTWARPESDGGSEIQOILEREKK 21697
QY 183 ----- 182
Db 21698 STRWVKVSKRPISETRFKVTGLTEGNEYEFHVMAENAGVGPASGISRLIKCREPVNPP 21757
QY 183 ----- 182
Db 21758 GPPTVVKVTDTSKTTVSLWSKPVFDGGMEIIGYIIECMCKTDLGDWHKVNAEACVKTRYT 21817
QY 183 ----- 182
Db 21818 VTDLOAGEYKFRVSAINGAGKGDSCVETGTIKAVDRLTAPELDIDANFKQTHVVVRAGAS 21877
QY 183 ----- 182
Db 21878 IRLFIAYQGRPTPTAVWSKPDNSILRADIHTTDSFSTLTVCNCRNDAGKYTLTVENNS 21937
QY 183 ----- 182
Db 21938 GSKSITFTVKVLDTGPGGPITFKDVTGRSATLMDAPLLDGGARITHHYVVEKREASRS 21997
QY 183 ----- 182
Db 21998 WQVISEKCTRQIFKVNDLAEGVPYFVRVSAVNEYGVGEPEYEMPEPIVATEQAPARRLDV 22057
QY 183 ----- 182
Db 22058 VDTSSSAVLAWLKPDBHGGSRITGYLLEMRQKGSGLWVEAGHTKOLTFTVRLVEKTEY 22117

Qy	171	-----	170
Db	17738	KPEHGGSRVLGYVVEMQPKTEKWSIVAESKVCNAVVTGLSSGOEYQFRVKAYNEKGKS	17797
Qy	171	-----	170
Db	17798	DPRLGVPEVIAKDLTIQPSLKLPTNTYSIQAGEDLKIEPIVIGRPRPNISWKQGEPLKQ	17857
Qy	171	-----	170
Db	17858	TTRVNVEETATSTVLHIKEGKNDDGKYTVTATNSAGTATENLSVIVLEKPGPPGVPYRF	17917
Qy	171	-----	170
Db	17918	DEVSADFVVISWEPPAVTGGCISNYIVEKRDTTTTTWHMVSATVARTTIKITLKTGTE	17977
Qy	171	-----	170
Db	17978	YQFRIFAENRYKSAPLDSKAVIYOYPEKPGPGTPEVTSISKQMDLVQWHPEVNDGGT	18037
Qy	171	-----	170
Db	18038	KIICGYHLBQEKENSLWVKLNKTIQDPRFKFTTGLDGLGEYEFKVSABNIYVIGIKPSKVS	18097
Qy	171	-----	170
Db	18098	ECFVARDPCDPPRPEAIVITRNNVTLKWKKPAYDGGSKITGYIVEKKDLPGRWMKASF	18157
Qy	171	-----	170
Db	18158	TNVLETEFTVSGLVEDQRYEFVFIARNAAGNFSEPSDSSGAITARDEIDAPNASLDPKYK	18217
Qy	171	-----	170
Db	18218	DVIVVHAGETVLEADIRGKPIPDVWMSKDKLEETAARMEIKSTIQKTTLVVKDCIRT	18277
Qy	171	-----	170
Db	18278	DGGYIILKSNVGGTKSIPITVKVLDPRGSPGPKLVGTVAEKCYLAWNPPLQDGGANI	18337
Qy	171	-----NKYTT-----	175
Db	18338	SHYIIKRETRSLSWTQVSTEQALNYKVTLLPGNEYIFRVMVANKYGIGEPLESGPVT	18397
Qy	176	-----	175
Db	18398	ACNPYKPGPSTPEVSATIKDSMVVTWARVPDGGTEIEGYILLEKRDKEGVRWTKCNKK	18457
Qy	176	-----	175
Db	18458	TLTDLRLRVTLGTEGHSYEFRAAENAAAGVGEPSVIFYRACDALYPPGPPSNPKVTDIT	18517
Qy	176	-----	175
Db	18518	SRSSVSLAWSKPIYDGGAPVKGYVVEVKEAAADEWTTCTPTGLQGKQFTVTKLKENTPEY	18577
Qy	176	-----	175
Db	18578	NFRICAINSEGVGEPATILPGSVAAQRIEPEIELDADLRKVVVLRASATLURLFTTKGR	18637
Qy	176	-----	175
Db	18638	PEPEVKWEKABGILTDRAQIEVTSFTMLVIDNTRFDSGRYNLTLNENSGSKTAFVNVR	18697
Qy	176	-----	175
Db	18698	VLDSPSAPVNLTIREVKKDSVTLSEWPEPLIDGAKITNYIVEKRETTKAYATITNNCTK	18757
Qy	176	-----	175
Db	18758	TTFRIENLQEGCSYFRVLASNEYIGLPAETTEPVPKVSPEPLPPGRVTLVDVTRNTATI	18817
Qy	176	-----	175

Db	18818	KWEPESDGGSKITGYVVMOTKGEKWKSTCTQVKLTLEATISGLTAGBEYVFRVAANKE	1887
Qy	176	-----	175
Db	18878	GRSDPRQLGVPVIARDIEIKPSVELPFFHFNVKAREQLKIDVPKGRPOATVNRKDGOT	18937
Qy	176	-----	175
Db	18938	LKETTRVNVSSKTYTSLSIKEAKEDYGVYELCVSNSAGSITVPITIIVLDRPGPGPI	18997
Qy	176	-----	175
Db	18998	RIDEVSCDSITISWNPEYDGGCISNIYVEKKETTSTTWHIVSOAVARTSIKIVRLTTG	19057
Qy	176	-----	175
Db	19058	SEYQFRCAENRYGKSSYSSAVVAEYPPFPBPBPQPKVHVHATKSTMLTWQVPVNDG	19117
Qy	176	-----	175
Db	19118	GSRVIGYHLEYKERSILWSKANKILADTQVKYSGLDGLMYEYRVYAENIAGIKCSK	19177
Qy	176	-----	175
Db	19178	SCEVPARDPCDPQGPQEVNTITRKSVLKWKSPHYDGGAKITGYIVERRELDPGRWLKC	19237
Qy	176	-----	175
Db	19238	NTNTOETYFEVTELTQRYEYFRVAFARNAADSVPSESTGPIIVKDDVPPRVMMDVK	19297
Qy	176	-----	175
Db	19298	FRDIVVKAGEVLKINADIAGRLPLVSWAKDGIIEERARTEIISTDNHTLLTVKDCIR	19357
Qy	176	-----	175
Db	19358	RDTOYVLTLNKNAVGRSAVNCKVLDKPGPPAGPLEINGLTAEKCSLSWGRPOEGGAD	19417
Qy	176	-----	175
Db	19418	IDYYHRKRETSHLAWTICEGELQWTSCKVTLLKGNEYIFRVTGVNKYGVGPLESVAI	19477
Qy	176	-----	175
Db	19478	KALDPFTVPSPPTSLEITSVKESMTLCWSRPESDGGSEISGYIEREKNSLRWVRNK	19537
Qy	176	-----	175
Db	19538	KPYDLRVKSTGLREGCEYERYVAENAAAGLSLPSETSPLIARAEDPVFLPSPSPKPIVD	19597
Qy	176	-----	175
Db	19598	SGKTTITIAWVKPLFDGGAPITGYTYEYKSDDTWKTSIOSLRGTEYISGLTTGAEYV	19657
Qy	176	-----	175
Db	19658	FRVSVNKVGASDPSSDPOIAKEREPEPLFIDISEMRKTLIVKAGASFMTVYFRGRP	19717
Qy	176	-----	175
Db	19718	VPNVLWSKPDTLRTRAYVDTDSRTSLT IENANRNDSGKYTLT IONVLSAASLTLVVKV	19777
Qy	176	-----	175
Db	19778	LDPGPPTNITVDVTKEAVLSMDVPENDGGA PVKNYHIEKREASKAWSVTNNCNRL	19837
Qy	176	-----	175
Db	19838	SYKVTNLQEGAIYFRVRSGENEFGVIPAETKEGVKITEKPSPEKLGVTISIKDSVSLT	19897
Qy	176	-----	175

Db 15518 LOEGCSYFRVLAENEYIGLPAETAESVKASERDLPQKITLMDVTRNSVLSWEKPEH 15577
QY 168 ----- 167
Db 15578 DGGSRILGYIVEMQTKGSDKWATCATVKTEATITGLIQGEYSFRVSAQNEKGISDPRQ 15637
QY 168 ----- 167
Db 15638 LSPVIAKDLVIPPAPKLLFNFTTFLAGEDLKVDVFFIGRPTTAVTWHKDNVPLKQTTVR 15697
QY 168 ----- 167
Db 15698 NAESTENNSLLTIKDACREDVGHYVVKLTNSAGEAIELNVILDKPGPPTGPVXMDVET 15757
QY 168 ----- 167
Db 15758 ADSITLSWGPYPKYDGGSSINNYIVEKRDITSTTTWQIVSATVARTTIKACRLKTGCEYQFR 15817
QY 168 ----- 167
Db 15818 IAAENRYGKSTYLNSEPTVAQYFPKVPVPGPTPVVTLSSRDSMEVQWNEPISDGGSRVIG 15877
QY 168 ----- 167
Db 15878 YHLERKERNISILWKLNTPIPTQTKFTGTLEEGVEYEFVSAENIVIGRPSKVSECV 15937
QY 168 ----- 167
Db 15938 ARDPCDPPGRPEALIVTRNSVTLOWKKPTYDGGSKITGYIVEKKELPEGRWMAKSFNTII 15997
QY 168 ----- 167
Db 15998 DTHFEVTGLVEDHRYEFVRVIAARNAAGVSEPSSTCAITARDEVDPPIRISMDPKYKDTIV 16057
QY 168 ----- 167
Db 16058 VHAGESFKVDADIYKPIPTIOWIKGDQELSNTARLEIKSTDFATSLSVKDAVRVDSGNV 16117
QY 168 ----- 167
Db 16118 ILKAKNVAGERSVTNVKVLDRPPEGPVVISGYTAEKCTLAWKPLQDGGSDIINIV 16177
QY 168 ----- 167
Db 16178 ERRETSRLVMTVVDANVOTLSCKVTKLLEGNEYTFRIMAVNKYGVGEPISEPVVAKNPF 16237
QY 168 ----- 167
Db 16238 VVPDAPKAPETVTTKDSMIVVWERPASDGGSEILGYVLEKRDKEGIRWTRCHKRLIGEL 16297
QY 168 ----- 167
Db 16298 RLURVTGLIENHDEYFRVSAENAGLSEPPSAYQKACDPIYKPGPPNPNKVIDITRSSV 16357
QY 168 ----- 167
Db 16358 FLWSKPIYDGCIEQGYIVEKCDVNWGEWMTCTPTTGINKTNIEVEKLEKHEYNFRIC 16417
QY 168 ----- 167
Db 16418 AINKAGVGEHADVPPIIVEEKLEAPDIDLDELKRIINIRAGGSLRLEFVPIKGRPTPEV 16477
QY 168 ----- 167
Db 16478 KWGKVGDEIRDAAIIDVTSSTSLVDNVNRYDSCKYTLTLENSSGTKSAFVTVRVRLDTP 16537
QY 168 ----- 167
Db 16538 SPPVNLKVEITKDSVITWEPPLDGGSKINNYIVEKREATRKSAAVVTNCHKNWKI 16597
QY 168 ----- 167
Db 16598 DQLQEGCSYFRVTAENEYIGLPAQTADPIKVAEVPQPGKITVDDVTRNSVLSWTKP 16657

QY 168 ----- 167
Db 16658 EHDGSKIIQYIVEMQAKHSEKSEKARVKSILQAVITNLTOGEEYLFVRVAVNEKGRSDP 16717
QY 168 ----- 167
Db 16718 RSLAVPIVAKDLVIEPDKVPAFSSYSVOQDLKMEVPISGRPKPTITWTKGLPKQTT 16777
QY 168 ----- 167
Db 16778 RINVTDSLDLTLTSIKETHKDDGGQGYITVANVWCQTASIEIVTLDRKDDPKPGVKFDD 16837
QY 168 ----- 167
Db 16838 VSAESITLSWNPPLYTGCOITNYIVOKRDTTTTVMVDVVSATVARTTLKVTCLKTGTETQ 16897
QY 168 ----- 167
Db 16898 FRIFAENRYGOSFALESDFIVAQYPKPEPGPTPFATAISKDSMVIQWHEPVNNGGSPV 16957
QY 168 ----- 167
Db 16958 IGYHLERKERNISILWKLNTKVNKTIHDTQFKAQNLBEGIEYEFVRYAENIVGVGKASKNSEC 17017
QY 168 ----- 167
Db 17018 YVARDPCDPPGTPEPIWVKRNEITLQWTKPVYDGGSMITGYIVEKRDLPDGRWMAKSFNT 17077
QY 168 ----- 167
Db 17078 VIETOFTVSGLTEDQRYEFVRVIAKNAAGAIKSPSDSTGPTAKDEVLPRIISMDPKFRDT 17137
QY 168 ----- 167
Db 17138 IYVNAVGETFRLEADVHGKPLPTIEWLRGDKEIEESARCEIKNTDFKALLIVKDAIRIDGG 17197
QY 168 ----- 167
Db 17198 QYILRASNVAGSKSPVNVKVLDRPPEGPVQVGTSEKSLTWSPPLQDGGSDISHY 17257
QY 168 ----- 167
Db 17258 VVEKRETSRLAWTVVASEVNTNSLKVTKLLEGNEYVFRIMAVNKYGVGEPISEAPVLMKN 17317
QY 168 ----- 167
Db 17318 PFVLPDPPKSLSEVTNIAKDSMTVCWNRDPDSDGGSEIIGYIVEKRDGRGIRWIKCNKRIT 17377
QY 168 ----- 167
Db 17378 DLRLRVTLGTEDHEYEFVRVSAENAGVGEPSPATVYKACDPVFKPGPPTNAHIVDTTKN 17437
QY 168 ----- 167
Db 17438 SITLAWKPIYDGSSEILGYVVEICKADEEEMQIVTPTQGLRVTRFEISKLEHOEYKIR 17497
QY 168 ----- 167
Db 17498 VCALNKVGLGEATSVPGTVKPEDKLEAPELDLDELKRIINIRAGGSLRLEFVPIKGRPMP 17557
QY 168 ----- 167
Db 17558 ETTWSREGEFTDKVQIEKGVNYTQLSIDNCNDRDAGKYILKLENSGSKSAFVTVKVLVD 17617
QY 168 ----- 170
| : |
Db 17618 TPGPPONLAVKEVRKDSAFLVWEPPIIDGAKVKVNIIDKRESTRKAYANVSSKCSKTSF 17677
QY 171 ----- 170
Db 17678 KVENLTEGAIYFRVMAENRFGVGVVETVDVAKAEPSPPGKVTLTLDVTSQTSASLMWE 17737

QY 161 ----- 160
Db 13358 SYFRVFAENYXGIGDPGETRDVAKASOTPGPVVDLKVRSVSKSSCSIGKKPHSDGSR 13417
QY 161 ----- 160
Db 13418 IIGYVVDFLTEENKQVRMKSLSIQSAKDLTEGKEYTFRVSAENEGECTPSEITVVAR 13477
QY 161 ----- 160
Db 13478 DDVVAPDLULGLPDLCLYLAKENSNFLKPIKPKPAPSVSWKKGEDPLATDTRVSESS 13537
QY 161 ----- 160
Db 13538 AVNTTLIVDCQSDACKYITILKNVAGTKEGTISIKVVKPGIPTGPIKDFEVTAEAMT 13597
QY 161 ----- 160
Db 13598 LKWAPPKDDGGSEITNYILEKRDVNNKWVTCASAVQKTFRTVRLHEGMEYTFRVSAEN 13657
QY 161 ----- 160
Db 13658 KYGVGGLKSEPIVARHPFDVPDAPPPNIVDRHDSVSLTWTDPKKTGGSPITGYHLEF 13717
QY 161 ----- 160
Db 13718 KERNSLLWKRAKTPIRMRFKVTYGLTEGLEYEFRVMAINLAGVGKPSLPSEPVVALDPI 13777
QY 161 ----- 160
Db 13778 DPPGKPEVINITRNSVTLIWTEPKYDGGHKLTYIVEKRDLPKSNWKAHVNVPECAFT 13837
QY 161 ----- 160
Db 13838 VTDLVEGGKEYFRIRAKNTAGAISAPSESTETIICDEYEAPTIVLDPTIKDGLTIKAGD 13897
QY 161 ----- 160
Db 13898 TIVLNAISILGKPLPKSSWSKAGDIRPSDITQITSTPTSSMLTIKYATRKDAGEYITIA 13957
QY 161 ----- 160
Db 13958 TNPFGTKVEHVKTIVLDVPGPGPVEISNVSAEKATLTWTPLEDGGSPIKSYILEKRET 14017
QY 161 ----- 160
Db 14018 SRLLTWVSEDIOCRHVATKLIQNEIYFRVSAVNHYGKEPVQSEPVKWDVRFPGPGP 14077
QY 161 ----- 160
Db 14078 PEKPEVSNVTKNATVSWKRPVDDGGSEITGYHVERREKSLRWVRAIKTPVSDLRCKVT 14137
QY 161 ----- 160
Db 14138 GLOEGSTYEFVSAENRAGIGPPSEASDVLMKDAAYPGPPSPNPHVTDTRKKSALWG 14197
QY 161 ----- 160
Db 14198 KPHYDGGLEITGYVVEHQVGDEAWIKDTGTALRITQFVVDLQTKKEYNFRISAINDA 14257
QY 161 ----- 160
Db 14258 GVGEPAVIPDVEIVEREMAPDFELDAELRRTLVRAGLSIRIFVPIKGRPAPEVWTWKN 14317
QY 161 ----- 160
Db 14318 INLKNRANIENTESFTLLIIEPCNRYDTGKFVMTIENPAGKSGFVNVRLDTPGVLNL 14377
QY 161 ----- 160
Db 14378 RPTDITKDSVTLHWDPLIDGGSRTINYIVEKREATRKSYSSTATTKCHKTKYKVTGLSEG 14437
QY 161 ----- 160

Db 14438 CEYFFRMAENYXGIGEPTTETTEPVKASEAPSPDLSLINDITKSTVSLAWPKPHDGG 14497
QY 161 ----- 160
Db 14498 KITGVIEAQRKGDQWTHITTVKGLECVVRLNLTGEGETTFQVMVNASGRAPRESRPV 14557
QY 161 ----- 160
Db 14558 IVKEQTMLPELDLURGIYQKLVIAKAGDNKVEIPVGRPKPTVTWKKGDOILKQTORVNF 14617
QY 161 ----- 160
Db 14618 ETTATSTILNINECVRSDSGPPLTARNIVGEVDVITIQVHDIPGPPTGPIKDFEVSSD 14677
QY 161 ----- 160
Db 14678 FVTFSWDPDPPDGGVPISNVVEMROTSTTWVELATTVIRTYYKATRLTTGLEYOFRVK 14737
QY 161 ----- 160
Db 14738 AQNRVGVGPGITSAWIVANYPVKVPGPGPTQVTAVTKDSMTISWHEPLSDGGSPILGYH 14797
QY 161 ----- 160
Db 14798 VERKERNGILQVTSKALVPGNIFKSSGLTDGIAEYFRVIAENMACKSKPKSEPMLAL 14857
QY 161 ----- 160
Db 14858 DPIDPPGKVPPLNITRHTVTLKWAPEYTGCGFKITSYIVEKRDLPNGRMLKANFSNILEN 14917
QY 161 ----- 160
Db 14918 EFTVSGLTEDAAYEFRVIAKNAAGAISSPSPSDAITCRDDVEAPKIKVDVKFKDVIK 14977
QY 161 -----EWESNGO----- 167
Db 14978 AGEAFRLADVSGRPPTMEWSKDGKELEGTAKLEIKIADFTNLVKNKDSRRDSGAYTL 15037
QY 168 ----- 167
Db 15038 TATNPGFAKHIFNVKVLDRPGPEGLAVTEVTSEKVLSPFPDODGAKIDHIVYQK 15097
QY 168 ----- 167
Db 15098 RETSRLAWTNVASEVQVTKLVTKLLKGNEYIFRVMVANKYGVGEPESEPVLA VNPYGP 15157
QY 168 ----- 167
Db 15158 POPPKNPEVTITIKDSMVVCMWGHPSDGGSEIINYIVERRRKAGORWIKCNKKTLTDLRY 15217
QY 168 ----- 167
Db 15218 KVSGLTEGHEYEFIRMAENAGISAPSPSPFYKACDVFVKPGPGPNRVLDTSRSSISI 15277
QY 168 ----- 167
Db 15278 ANKKPIYDGGSEITGYVWEIALPEDEWQIVTPPAGLKAATSYITGLTENQEKIRIYAM 15337
QY 168 ----- 167
Db 15338 NSEGLGEPALVPCTPKAEDRMLPPEIELDADLRKVVTIRACCTLRLFVPIKGRPDPEVKW 15397
QY 168 ----- 167
Db 15398 ARDHGESLDKASIESASSYTLIVGNVNRFDGSKYILTVENSNGSKSAFVNVRLDTPGP 15457
QY 168 ----- 167
Db 15458 PODLKVKVTKSVTLTWDPPLLDGGSKIKNYIVEKRESTRKAYSTVATNCHKTSWKVDQ 15517
QY 168 ----- 167

Db 11138 PVINYIVEKQDTRKDTGWVSSGSSKTKLIPHLOKGCYVFRVRAENKIGVGPLDSTP 11197
QY 154 ----- 153
Db 11198 TVAKHKFSPSPGKPVVTDITENAAVSWTLPKSDGSGPITGYMERREVTGKWVRNK 11257
QY 154 ----- 153
Db 11258 TPIADLKFRVTGLYEGNTYEFVRFAENLAGLSKPSSPDIKACRPKPPGPPINPKLKD 11317
QY 154 ----- 153
Db 11318 KSRETADLVWTKPLSDGSGPILGYVVECOKPGTAQWNRINKDELIRQCAFVPGLIEGNE 11377
QY 154 ----- 153
Db 11378 YRFRIKAANIVGEGEPRELAESVIAKDILHPPEVELDVTCDVITVRVGTIRILARVGK 11437
QY 154 ----- 153
Db 11438 RPEPDITWTKGKVLVREKRVLDIOLPRVELQIKEAVRADHGKYIISAKSSGHAQGSA 11497
QY 154 ----- 153
Db 11498 IVNVLDROPQCONLKVTNTKENTISWENPLDNGGSEITNFIVEYRKNQKGSIVASD 11557
QY 154 ----- 153
Db 11558 VTKRLIKANLANNEYFRVCAENKVGVTIETKTPIILAINIDRPGEPENLHIADCK 11617
QY 154 ----- 153
Db 11618 TFYILKWRPRPDYDGGSPNLSYHVERRLKGGDDWERYHKGSIKETHYMDRCVENQIYEF 11677
QY 154 ----- 153
Db 11678 VQTKNEGGSWVKEEVVVKEDLQKPVLDLKLSGVLTWKAGDTRLEAGVRGKPPPEVA 11737
QY 154 ----- 153
Db 11738 WTKDKDATLITRSPRVKIDTRADSSKSLTKAKRSDGKVVYVATNTAGSFVAYATVNVL 11797
QY 154 ----- 153
Db 11798 DKPGVVRNLKIVDVSSDRCTVCDPDDGCGEIQNYILEKGETKRWVWSTYSATVLTGP 11857
QY 154 ----- 153
Db 11858 TTVTRLIEGNEYIFRVAENKIGTGPTBSKPVIAKTKYDKPGRPDPEVTKVSKEMTV 11917
QY 154 ----- 153
Db 11918 VNNPEYDGGKSTGYFLEKKEKHSRWVPVNSAIPERRMKVQNLPLPDHEYQFRVKAEN 11977
QY 154 ----- 153
Db 11978 EIGIGEPSLSPRVVAKDIEPPGPPTNFRVVDTTKHSITLWGKPVYDGGAPIIIGYVVE 12037
QY 154 ----- 153
Db 12038 MRPKIADSPDEGWKRCNAAQLVRKEFTVTSLDENQEYFRVCAQNVGIGRPAELKEA 12097
QY 154 ----- 153
Db 12098 IKPKEILEPPEIDLASMRKLVIRAGCPIRLFAIVRGPAPKVTRKVGIDNVVRKGOV 12157
QY 154 ----- 153
Db 12158 DLVDTMAFLVIPNSTRDDSGKYSLTLPNAGEKAVFNVVRVLDTPGVSDLKVSDVTKTS 12217
QY 154 ----- 153
Db 12218 CHVSWAPPENDGGSQVTHYIVEKREADRKTSTVTPVEVKKTSFHVNLNLPVPGNEYVFRVTA 12277

QY 154 ----- 153
Db 12278 VNEYGPGVPTDVPKPVLASDPLSEDPDPKLEATEMTKNSATLAWLPPLRDGGAIDGVI 12337
QY 154 ----- 153
Db 12338 ISYREEQPADRWTEYSVVVKDLSLVVTGLKEGKYYKFRVAARNAVGSUPREAEGYEAK 12397
QY 154 ----- 153
Db 12398 EQLLPKILMPEQITIKAGKKLRIEAAHVYKPHPTCKWKKGDEVVVTSSHAVHRAADSS 12457
QY 154 ----- 153
Db 12458 ILI IKDVTBKDSGYSLTAENSSGTDOKIKVVMADAPGPPPPFDISDADACSLSMH 12517
QY 154 ----- 153
Db 12518 IPLEDGGSNITNYIVEKCDVSRGDWWTALASVTKTSCRVGKLIPGOEYIFRVAENRFGI 12577
QY 154 ----- 153
Db 12578 SEPLTSPKVAQFPFGVPSEPKNARVTKVKNKDCIFVWDRPDSGSGPIIGYLIERKERN 12637
QY 154 ----- 153
Db 12638 SLLWVKANDTLVRSTEYPCAGLVEGLEYSFRIVALNKAAGSSPKPTEYVTARMPVDPGP 12697
QY 154 ----- 153
Db 12698 KPEVIDVTKSTVLIWARKPHDGGSKIIGYFVEACKLPGDKWVRCNTAPHQIPQEEYTAT 12757
QY 154 ----- 153
Db 12758 GLEBRAQYQFARAIARTAVNISPSESDPVTILAENVPDRIDLSVAMKSLITVKAGTNVC 12817
QY 154 ----- 153
Db 12818 LDATVFGKPMPTVSMKKDGTLLKPABGKMAQORNLCTLELFSVNRKSDGYTITAENSS 12877
QY 154 ----- 153
Db 12878 GSKSATIKLVLDKFGPPASVKINKMYSDRAMLSWEPPLDGGSEITNYIVDKRETSRPN 12937
QY 154 ----- 153
Db 12938 WAQVSATVPITSCSVEKLLIEGHEYQPRICAENKYGVGDPVFTPEPAIAKNPYDPPGRCDPP 12997
QY 154 ----- 153
Db 12998 VISNITKDHMTYSWKPPADDGSGSPITGYLLEKRETOAVNWKVNRKPIIERTLKATGLOE 13057
QY 154 ----- 153
Db 13058 GTEYEFRTAINKAGPKPSDASKAAYARDPOYPAPPAPKVIYDTRRSVLSWKGKPAY 13117
QY 154 ----- 153
Db 13118 DGGSPIIGYLVEVKRADSDNNWVRNLPQNLQKTRFEVTGLMEDTQYQFRVYAVNKGISD 13177
QY 154 ----- 160
Db 13178 PSDVPDKHYPKDILLIPPEGEHDADLRKTLILRAGVTMRLYVPVKGRPPPKITWSKPNVNL 13237
QY 161 ----- 160
Db 13238 RDRIGLDIKSTDFTPLRCENVNKYDAGKYILTLNESCCKEYTIIVVKVLDTPGPPINVT 13297
QY 161 ----- 160
Db 13298 VKEISKDSAYTWEPPIIDGGSPIINYYVQKRDARERKSWSTVTECKSTSRFVNPNEEGK 13357

QY 133 ----- 132
Db 8978 YTVASNRLGVSFRNVHVEYDRPSPRNLA VTDIKAESCYLTWDAPLNDGSGSEITHYVI 9037
QY 133 ----- 132
Db 9038 DKRDASRKAEBEVTNTAVEKRYGIWKLPNGQYEFVRVANKYGISDECKDKVVIQD 9097
QY 133 ----- 132
Db 9098 PYRLPGPGPKVLARTKGSMLVSWTPPLDNGGSPITGYWLEKREGEPSYWSRVSRAPIT 9157
QY 133 ----- 132
Db 9158 KVGLKGVEFNVPRLLGKVKQFRAMAINAAGIGPPSEPSDEVAGDPIFPFGPPSCPEVK 9217
QY 133 -----PPSRD----- 137
| : : :
Db 9218 DTKSSISLGWKPPAKDGGSPKIGYIVEMOEGTDMKRVNEPDKLITTCCECVVNLKEL 9277
QY 138 ----- 137
Db 9278 RYFRVRKAVNEAGESEPSDTTGEIPATDIQEBPEVFIDIGAQDCLVCRAKAGSIRIPAVI 9337
QY 138 ----- 137
Db 9338 KGRPTPKSSNEFDGAKAMKMGVHDIPEDAQIETAENSSVILIIPECKRSHTGKYSITAK 9397
QY 138 ----- 137
Db 9398 NKAGQKTANCRVKVMDVPGPKDLKVSIDITRGSCRLSWKMPDDGGDRIKGYVIEKRTID 9457
QY 138 ----- 137
Db 9458 GKAWTKVNPCGGSTTFVVPDLLSEQOYFFRVRANRFGIPGPVETIORTTARDPIYPPDP 9517
QY 138 ----- 137
Db 9518 PIKLGITLITKNTVHLSWPKPKNDGGSPTVHYIVECLANDPTGTKEANRQCNKRDVEEL 9577
QY 138 ----- 137
Db 9578 QFTVEDLVEGGEYFRVKAVNAAGSVKPSATVGCQCRDMPMSIDLKEFMEVEEGTNV 9637
QY 138 ----- 137
Db 9638 NIYAKIKGVFPPLTWFKAPPKKPDNKEPVLVYDTHVNKLVDVDTCTLVIPSRRSDTGLY 9697
QY 138 ----- 137
Db 9698 TITAVNLTGASKEMRLNVLGRPGPPVGPIKFESVSADQMTLSWFPKDDGSKIITNYVI 9757
QY 138 ----- 137
Db 9758 EKREANRKTWVHVSSEPKECTYITPKLLEGHEYVFRIMAQNKYIGEPDSEPETARNLF 9817
QY 138 ----- 137
Db 9818 SVPGAPDKPTVSSVTRNSMTVNWEEPEYDGGSPVTGYWLEMKDTSKRWKRVNRDPTIKAM 9877
QY 138 ----- 137
Db 9878 TLGVSYKVTGLIEGSDYQFRVYAINAAGVPASLPSPDPATARDPIAPPGPPFPKVTDWTK 9937
QY 138 ----- 137
Db 9938 SSADLEWSPLKDGSKVTGYIYEYKEGKEWEKGRKEVGTCLVVTGLKEGAFYKFR 9997
QY 138 ----- 137
Db 9998 VSAVNIAGICEPEVTDVEMKDRLVSPDLQDASVRDRIVVHAGGVIRIIAIVSGKPPP 10057
QY 138 ----- 137

Db 10058 TVTWNMNERLPOEATIETTATISSMVIKNCORSHOGVSYLLAKNEAGERKKTIIVDVLD 10117
QY 138 ----- 137
Db 10118 VPGVGTGPF LAHLNLTNESCCLTWFSPEDDGGSPITNYVIEKRESDRRAWTPVYTYVTRQN 10177
QY 138 -----ELTKNOVS 145
| : : :
Db 10178 ATVQGLIOGKAYFFRIAAENSICGPFVETSEALVIREPITVPERPEDLEVKEVTNKTVT 10237
QY 146 LT----- 147
| : : :
Db 10238 LTWNPPKYDGGSEIINVLESRLIGTEKPHKVTNDNLLSRKYTVKGLKEGDTYEYRSVAV 10297
QY 148 ----- 147
Db 10298 NIVGQKPSFCTKPIITCKDELAPPTLHLDPRDKLTIRVGEAFALTGRYSGKPKVSWFK 10357
QY 148 -----CLV-----KGF----- 153
| : : :
Db 10358 DEADVLEDDRRTHIKTTPATLALAEKIKAKRSDSGKYCVVVENSTGSRKGFQCVNVVDHGP 10417
QY 154 ----- 153
Db 10418 PVGVSFDEVTKDYMWISNKPPLDDGSGKITNYVIEKKEVGDVWMPVTSASAKTCKVS 10477
QY 154 ----- 153
Db 10478 KLEKGDYIFRIHAENLYGISDPLVSDSMKAKDRFRVPDAPQPIVTEVTKDSALVTWK 10537
QY 154 ----- 153
Db 10538 PHDGGKPIITNYILEKRETMKSRWARTKDIHPYTKFRVPDLPDLEGCOYEPRVSAENEIGI 10597
QY 154 ----- 153
Db 10598 GDPSPSPKPVFAKPIAKPPSPVYNPEALDITCNSVDLTWQPPRRHDGSGKITGYVIEYQKV 10657
QY 154 ----- 153
Db 10658 GDEWRANHTPESCPETRYKVTGLRDGQTYKFRVLAVNAAGESDPAHVPEPVLVKDRLE 10717
QY 154 ----- 153
Db 10718 PPELIILANWAREQHIVKVGDTLRLSALIKGVPPKVTWKEDRDAPTAKRIDVTPVGSKL 10777
QY 154 ----- 153
Db 10778 EIRNAAHEDGGIYSLTAVENPAGSKTVSVKVLVDKPGPPRDLEVSEIRKDCYLTKWEPL 10837
QY 154 ----- 153
Db 10838 DGGSVITNYVVERRDVSAQNSPLSATSKKSHFAKHLNEGNOYLFRVAENQYGRGPF 10897
QY 154 ----- 153
Db 10898 VETPKPIKALDPLHPGPKDLHHVDVDTESVSLVWNKPDRCGSPITGYLVEYQEGTQ 10957
QY 154 ----- 153
Db 10958 DWIKFKTVTNLECVVTGLOQKTYRFRVKAENIVGLGLPDTTPIIECQEKLVPPSVELDV 11017
QY 154 ----- 153
Db 11018 KLIEGLVWVAGTTVRPPAIIRGVPTAKWTTDGSSEIKTDEHYTVETDNFSSVLTKNCL 11077
QY 154 ----- 153
Db 11078 RRDTEYQITVSNAGSKTVAVHLTVLDVPGPPTGPIINILDVTPHEMTISWQPPKDDGGS 11137
QY 154 ----- 153

Db	6758	VVEGQYSFRVRAQNRIGVGKPSAATPEVKVADPIERPSPVNLTSDDQTOSSVOLKWE	6817
QY	133	-----	132
Db	6818	PLKGGSPILGYIIERCEEKDNWIRCNMKLVPELTQVTKGLERGNKYLVRSAENKAGV	6877
QY	133	-----	132
Db	6878	SDPSEILGLTADDAFVEPTMDLSAFKDGLEIVIPNPITILVPSTGYPRPTATWCFGDKV	6937
QY	133	-----	132
Db	6938	LETGDRVMKTLTSAELVISPERSDKGIYTLKLENRVKTIISGEIDVNVIAPSAPKEL	6997
QY	133	-----	132
Db	6998	KFGDITKDSVHLTWEPDDGGSPLTGYVVEKREVSRTWTVMDFVTDLEFTVPDLVQG	7057
QY	133	-----	132
Db	7058	KEYLFKVCARNKCGEPAYVDEPVNMSTPATVPDPENVKWRDRTANSIFLTWDPKND	7117
QY	133	-----	132
Db	7118	GGSHIKGYIVERCPRGSDKWVACGEPVETKMEVTGLEGKWYAYRVKTLNRQASKPSR	7177
QY	133	-----	132
Db	7178	PTEEIOAVDQEAPEIFDLVKKLAGLTVKAGTKIELPATVTVGKPEPKITWTKADMILKOD	7237
QY	133	-----	132
Db	7238	KRITIENVPKKSTVTIVDSKRSDTGTIYIEAVNVCGRATAVEVNVLDKPGPPAAFDITD	7297
QY	133	-----	132
Db	7298	VTNESCLLTWNPRDDGGSKITNYVVERRATDSEVWHKLSSTVKDTNFKATKLIPKEYI	7357
QY	133	-----	132
Db	7358	FRVAENWYGABPVQASPIYAKYQDPDPPGPPTRLEPSDITKDAVLTWCEPDGGSPI	7417
QY	133	-----	132
Db	7418	TGYWVERLDPTDKWRCRNMKPVKDTYRVKGLTNKKYRFRVLAENLAGPGKPSKSTEP	7477
QY	133	-----	132
Db	7478	ILIKDPIIDPPWPGKPTVKDVGKTSVRLAWTKPEHDGAKIESYVIEMLKTGTDEWVRVA	7537
QY	133	-----	132
Db	7538	EGVPTTQHLLPGLMEQGEYSFRVRAVNKAGESEPSDPSDVLCREKLYPPSPRWLEVIN	7597
QY	133	-----	132
Db	7598	ITKNTADLKWTVPEKDGSPITNYIVEKRDVRKGMQVDTTVTKDTKCTVTPLTEGSLIV	7657
QY	133	-----	132
Db	7658	FRVAENAIQSDYTEIEDSVLAKDTFTTGPYPYALAVVDVTKRHVDLKWPEPKNDGGRP	7717
QY	133	-----	132
Db	7718	IQRVYIEKKERLGRWVKAGTAGPCNFRVTDVIEGTVEQFOVRAENAGVGHPSSEPT	7777
QY	133	-----	132
Db	7778	ILSIEDPTSPSPPLDLHVTAGRKHIAIAWKPPKNGGSPILGYHVEMCPVGTCKWVRV	7837
QY	133	-----	132
Db	7838	NSRPIKDLKFKVEGVVPOKEYVLRVRAVNAIGVSEPSSEISNVVAKDPDCKPTIDLETH	7897

QY	133	-----	132
Db	7898	DIIVIEGEKLSIPVFFRAVPVPTVSWHKDGKEVKASDRLTMKNDHISAHLEVPKSVRADA	7957
QY	133	-----	132
Db	7958	GIYITITLENKLSATASINVKVIGLPGCKDIKASDITKSSCKLTWBPPEFDGGTPIIHY	8017
QY	133	-----	132
Db	8018	VLEREAGRRTYIPVMSGENKLSWTVKDLIPNGEYFFRVRANKVVGGEYIELKNPVAQ	8077
QY	133	-----	132
Db	8078	DPKOPDPDDVEVHNPTAEAMTITWKPPPLYDGGKIMGYIIIEKIAKEERWKRNEHLV	8137
QY	133	-----	132
Db	8138	PILTYTAKGLEEGKEYQFRVRAENAGISEPSRATPPTKAVDPIDAPKVILRTSLEVRG	8197
QY	133	-----	132
Db	8198	DEIATDASISGSPYPTITWIKDENVIYPEIKKRAAPLVRRRKGEVQEEFFVLPLTQRL	8257
QY	133	-----	132
Db	8258	SIDSKKESQLRVDRSLRDPDHGLYMIKVENDHGIAKAPCTSVLDTPGPPINFVEDIR	8317
QY	133	-----	132
Db	8318	KTSVLCKWEPDLDGGSEIINYTLKKDKTKPDSEWIVTSTLRHCKYSVTKLIEGKEYL	8377
QY	133	-----	132
Db	8378	FRVRAENRFGPPCVSKPLVAKDPGPPDAPDKPIVEDVTSNSMLVKNWNEPKDNGSPIL	8437
QY	133	-----	132
Db	8438	GYWLEKREVNSTHWSRVNKSLLNALKANVDGLLEGLTYFRVCAENAGPGKFPSPDPK	8497
QY	133	-----	132
Db	8498	TAHDPISPPGPPIPRVTDTSSTTIELEWEPAPFNGGGEIVGYFVDKQLVGTNKNWCRTEK	8557
QY	133	-----	132
Db	8558	MIKROYTVKEIREGADYKLRVSAVNAAGBPPGETQPTVAEPQEPAPVELDVSVKGGI	8617
QY	133	-----	132
Db	8618	QIMAGKTLRIPAVVTGRPVPTKVWTKEEGELDKDRVIDNVGTKSELLIKDALRKDHGRY	8677
QY	133	-----	132
Db	8678	VITATNSCGSKFAAARVEVDVPGPVLDLKPVVTVNRKMLLNWSDPEDDGGSEITGFIE	8737
QY	133	-----	132
Db	8738	RXDAKMTWRQPIETERSKCDITGLLEGOEYKFRVIAKNKFGCGPPVEIGPILAVDPLGP	8797
QY	133	-----	132
Db	8798	PTSPERLTYTERQSTITLDWKEPRNGGSPIQGYIIIEKRRHDKPDFERVNKRCLCPTTSF	8857
QY	133	-----	132
Db	8858	LIVENLDEHOMYEFVRKANVEIGESEPSLPNVVQDDEVPPTIKLRLSVRGDTIKVKAGE	8917
QY	133	-----	132
Db	8918	PVHIIPADVTGLPWPKIENSKNETVIEKPTDALQITKEEVSRSSEAKTELSIPKAVREDKCT	8977

Qy	115	-----	114
Db	4598	SGGKPPDEAPFTYQKAPLKFVKKEIKDIILTESEFVGSSAIFECVSPSTAITWMKO	4657
Qy	115	-----	114
Db	4658	GSNRESKPHRFIADGKORKLHIIDVQUSDAGEYTCVLRGNKEKTSTAKLVVEELPVRF	4717
Qy	115	-KTIS- :	118
Db	4718	VKTLEEEVTVVKQPLYLSCELNKERDVVRKDKIVVEKPGRIVPGVIGLMRALTINDA	4777
Qy	119	-----	118
Db	4778	DDTAGTYTVVENANNLECCSVKVEVIRDLWPKPIRQHVKPGTAIFAACDIAKDTP	4837
Qy	119	-----	118
Db	4838	NIKFKGYDEIPAEPNDKTEILRDGNHLYLKIKNAMPEDIAEYAVEIEGKRYPAKLTLGE	4897
Qy	119	-----	118
Db	4898	REVELLKPIEDVTIYEKESADAEISEADIPGQWKLKCELLRPSPTCEIKAEKGKRLT	4957
Qy	119	-----	118
Db	4958	LHKVKDQAGEVLYQALNAITTAILTVKEIELDFAVPLKDVTPERRQARFECVLTREAN	5017
Qy	119	-----	118
Db	5018	VIWSKGPDIIKSDKFDIIADGKKHILVINDSQDFDEGVYTAEEVGKTSARLFTVGTGIRL	5077
Qy	119	-----	118
Db	5078	KFMSPLEDOTVKEGETATFVCELSHEKMHVWVFNDAKLHTRSIVLISSEKTHKLENKE	5137
Qy	119	-----	118
Db	5138	VTLDDISQIAQVKELSTAQLKVLADPYFTVKLHDKTAVEKDEITLKCEVSKDVPVKW	5197
Qy	119	-----	118
Db	5198	KDGEIIVPSPKYSIKADGLRILKIKADLKRGVEYVDCGDTKTKANVTVEARLIEVE	5257
Qy	119	-----KAGOP- 	124
Db	5258	KPLYGVEVGETAHFEIELSEPDVHGQWKLKGQPLTASPCDEIIEDGKKHILILHNCQL	5317
Qy	125	-----REPO- :	128
Db	5318	GMTGEVSFOAANAKSAANLKVKELPLIFITPLSDVKVFKEDEAKFECEVSREPKTFRWLK	5377
Qy	129	-----	128
Db	5378	GTQBITGDRPELIKDGTKHSNVIKSAFAFEDEAKYMEAEADKHTSGKLIIEGIRLKFLTP	5437
Qy	129	-----	128
Db	5438	LKDVTAKEKESAVTVVELSHDNIRVKWFKNDQRLHTRTSVSMQDEGKTHSTTFKDLSDID	5497
Qy	129	-----	128
Db	5498	TSQIRVEAMGMSSEAKTLVLEGDPFYFTGKLQDYGVEKDEVILQCEISKADAPVKWFKDG	5557
Qy	129	-----	128
Db	5558	KEIKPSKNAVIKTDGKKRMLILKALKSDIGQYTCDCGTDKTSKGLDIEDREIKLVRPLH	5617
Qy	129	-----	128
Db	5618	SVEVNETARFETEISEDDIHANWKLKGEALLQTPDCEIKEEKGIIHSLVHLNCRLODTG	5677
Qy	129	-----	128

Db	5678	GVDFQAAVVKSSAHLRVKPRVIGLLRPLKDVTVTAGEATATEDCELSYEDIPVWYLKGGK	5737
Qy	129	-----	128
Db	5738	LEPSDKVVPRSEGVHTLTLRDVKLEDAGEVOLTAQDKPHTHANLEFVKPPVEFTKPLEDQ	5797
Qy	129	-----	128
Db	5798	TVEBEGATAVLECEVSRENAKVWFKNGTEILKSKYEIVADGRVRKLVHDCPTPEDIKTY	5857
Qy	129	-----	128
Db	5858	TCDARFKTSCNLNVPPHVEFLRPLTDLQVREKEMARFECELSRENAKVKWFKDGAIEIK	5917
Qy	129	-----	128
Db	5918	KGKYYDIISKGAIRILVINKLLODEAEYSCEVRTARTSGMLTVLEESAVFTKNLANIEV	5977
Qy	129	-----	128
Db	5978	SETDTIKLVCEVSKPGAIEVHWKGDEEIIETGRYEIILTEGKRILVIONAHLEDAGNYNC	6037
Qy	129	-----	128
Db	6038	RLPSSRTDGVKVHLEAAEFISKQNLLEGEKAFAEVCISIKESFPVQWKRDDKTLESG	6097
Qy	129	-----	128
Db	6098	DKYDVIADGKKRVLVVKDATLQDMGTVMVVGAAARAAHLTVIEKLRIVVPLKDTVRKEQ	6157
Qy	129	-----VYTL- 	132
Db	6158	QEVVFNCEVNTGAKAKWFRNEBAIFDSSKYIILOKDLVYTLRIRDAHLDDOQANYVSLT	6217
Qy	133	-----	132
Db	6218	NHRGENVKSANLIVEEEDLRIVEPLKDIEETMEKKSVTFWCKVNLNVTLWTKNGBEVP	6277
Qy	133	-----	132
Db	6278	FDNRSYRVYDKYKHLMTIKDCGFPDEGEYIVTAGQDKSVAELLITEAPTEFVEHLEDQTV	6337
Qy	133	-----	132
Db	6338	TEPDDAVFSQOLSREKANVWYRNGREIKEGKYKFEKDGSIHRLIINKDCRLDDECEYAC	6397
Qy	133	-----	132
Db	6398	GVEDKSRARLFVEEIPVEIIRPPQDILEAPGADVFLAELNKKDKEVQWLRNNWVVQG	6457
Qy	133	-----	132
Db	6458	DKHOMSEGIHRLQICDIKIPROGEYRFIAKDKEARAKLEAAAPKIKTADQLVDVVG	6517
Qy	133	-----	132
Db	6518	KPLTMVVPYDAYKAAEWFKEPENEPLSTKTTIDTTABOTSFRILEAKKGDGKRYKIVLONK	6577
Qy	133	-----	132
Db	6578	HGKAEGFINLKVIDVPGVVRNLEVTEFTDGEVSLAWBEPLTDGGSKIIGYVVERRDKRK	6637
Qy	133	-----	132
Db	6638	TWVLATDRASCEFTVTGLOKGGVEYLFVRSARNRVGTGEPVETDNPVEARSKYDVPGBP	6697
Qy	133	-----	132
Db	6698	LNVITIDVNRFGVSLTWBPPEYDGGAEITNVYIELROKTSIRWDTAMTVRAEDLSATVTD	6757
Qy	133	-----	132

Db 2378 NYSFTIPALGLSTGRVSYSVDVITPLKDNVNIIEGTAKVLECKVSPDVTSVKWLINDE 2437
QY 74 ----- 73
Db 2438 QIKPDRQAVKGTQKRLVINRTHASDEGPKLIVGRVETNCNLSVEKIKIIRGLRDLT 2497
QY 74 ----- 73
Db 2498 CTETQNVVEVELSHGIDVLNFKDKKEIKPSSKYKIEAHGKIYKLTVLNMMKDDGKYT 2557
QY 74 ----- 73
Db 2558 FYAGENWTSGLTVAGGAISPLTDQTVAESQEAUFCEVEANPDGSKGWLKHLPLTN 2617
QY 74 ----- 73
Db 2618 NIRSDEGHKRLIIAATKLDLDIGEYTYKVATSKTSAKLKVEAVKIKTKLNLTVETQD 2677
QY 74 ----- 73
Db 2678 AVFTVELTHPNVKGWIKNGVLESNEKYAISVKGTIYSLRIKNCIAIVDESIVYGRFGR 2737
QY 74 ----- 73
Db 2738 LGASARLHVETVKIITKKPKDVTALENATVAFEVSVSHDTPVKWPHKSVETKPSDKHRLV 2797
QY 74 ----- 73
Db 2798 SERKVHKLMLQNLSPSDAGEYTAVVQLECKAKLFEVLHITKTMKNIIEVPETKTAFEC 2857
QY 74 ----- 73
Db 2858 EVSHFNVPMSMLKNGVEIEMSEKIVVQGLHQLIIMNTSTEDSAEYTFVCGNDQVSAT 2917
QY 74 ----- 73
Db 2918 LTVTPIMTSMKLDINAEKDTITFEVTVNVEGYSYKWLKNGVEIKSTDKCOMTKKLTTH 2977
QY 74 ----- 73
Db 2978 SLNIRNVHFGDAADYTFVAGKATSTATLYVEARHIEFRKHIDKIVLEKKRAMFECEYSE 3037
QY 74 ----- 73
Db 3038 PDITVQWKKDDQLOITDRIKIOKEKYVHRLIIPSTRSDAGKTYTVVAGNVSTAKLVE 3097
QY 74 ----- 73
Db 3098 GRDVRIRSIKKEVQVIEKQRAVVEFEVNEDDVDHWHYKDGIEINPQVQERHXYVVERRIH 3157
QY 74 ----- 73
Db 3158 RMFISETRQSDAGEYTFVAGNRSSVTLYVNAPEPPQVLQELQPVTVQSGKPARFCAMIS 3217
QY 74 ----- 75
Db 3218 GRPQKISWYKEEQLLSTGFKCKFLHDGOEYTLLEAFEDAAVYTCENKNDYGVATTS 3277
QY 76 ----- 75
Db 3278 ASLSVEVPEVSPQEMPVYPPAIITPLQDVTVTSBQPARFCQVSGTDLKVSWSKDKK 3337
QY 76 ----- 83
Db 3338 IKPSRFRMTQFDYQLEIAEAYPEDEGTTFVANNVAGQVSSSTANLSLEAPESILHER 3397
QY 84 ----- 83
Db 3398 IEQEIEMEMKEFSSFLSAEEGLHSAELQLSKINETLELLESPPYPTKDFSEKEGTGP 3457
QY 84 ----- 87
Db 3458 IFIKEVSNADISMGDVATLSVTVIGIPKPIOWFNGVLLTPSADYKFVFDGDDHSLIIL 3517

QY 88 ----- 87
Db 3518 FTKLEDEGEYTCMASNDYKKTICSAYLKNKSGEGHKDTETESAVAKSILEKLGPCPHF 3577
QY 88 ----- 91
Db 3578 LKELPIRCAOGLPAIFEYTVVGEPAFTVTFKFNKQLCTSVYYTIIHNPNCSGTIVND 3637
QY 92 ----- 91
Db 3638 POREDSGLYICKAENMLGESTCAABELLVLLDTPCKAKSTPEAPEPQTPPLKGP 3697
QY 92 ----- 91
Db 3698 AVEALDSEIATFVKDTILKAALITEENQQLSYEHIKANELSSOLPLGAQELQILBQ 3757
QY 92 ----- 91
Db 3758 DKLTPESTREFLICINGSIHFOPLKEPSPNLQIIVOSQKTFSEKIGILMPEEPQAVLSD 3817
QY 92 ----- 91
Db 3818 TEKIPPSAMSTEQINSLTVEPLKTLAEPEGNYPOSSIEPPMHSHYLTLSVAEVLSLKEKT 3877
QY 92 ----- 91
Db 3878 VSDTNREQRVTLQKQEAQSALILSQSLAEGHVESLQSPDVMISQVNYEPLVPSEHSETEG 3937
QY 92 ----- 91
Db 3938 GKILIESANPLENAGDSAVRIEGLSLRFLPALEEKQVLLKEHSDNVMPDQIIESK 3997
QY 92 ----- 95
Db 3998 REPVAIKKQVQVQGRDILLKESLLSGIPEEQLNLKIQICRALQAAVASEQGLFSEWL 4057
QY 96 ----- 95
Db 4058 NIEKVEEAVNITQPRHIMCMYLVTSKSVTEVTIIIEDVDPQAMKLMELDALCAI 4117
QY 96 ----- 95
Db 4118 IYEEIDILTAEGRIOQAKTSLQEEMDSFSGSQKVEPITEPEVESKYLISTEEVSFNV 4177
QY 96 ----- 95
Db 4178 QSRVKYLDATPVTKGVASAVSDEKQDESLEKPSKEEKSSESGTEEVATVKIQAEGL 4237
QY 96 ----- 95
Db 4238 IKEDGPMIHTPLVDTPVSEGDIVHLTTSITNAKEVNWYFENKLVPSDEKFKCLODQNTYT 4297
QY 96 ----- 98
Db 4298 LVIDKVNTEHDHGEYVCEALNDSGKATSAKLTVVKRAAPVYKRIEPLVLEALHAKFT 4357
QY 99 ----- 106
Db 4358 CEIQSAPNVRFQWFKAGRIEYEDSKSIRSSKYISSLEILRTQVVDCGEYTCCKASNEYS 4417
QY 107 ----- 106
Db 4418 VSCTATLTVPVGEKKVKRLLPERKBPKEEVLKSVLRKRPEEPRKPEKLEKVKK 4477
QY 107 ----- 106
Db 4478 PAVPEPPPPKPEVEEVEVTVTKRERKIPEPTKVPEIKPAIPLPAPKPKPAEVKTKP 4537
QY 107 ----- 114
Db 4538 PPVEPTPIAAPVTVPVVGKKAEPKAAKPKPIKGVPKKTPSPIEAERKLRPG 4597

QY 10 ----- 9
Db 218 QTRIEKIEAHFDARSATVEMVIDGAAGQOLPHKTPRIPPKPKSRSPPTPSIAAKAQL 277
QY 10 ----- 9
Db 278 AROQSPSPIRHSPVRHVRAPTPSPVRSVSPAARISTSPRSPLLMRKTOASTVAT 337
QY 10 ----- 9
Db 338 GPEVPPKQEGYVASSSEAEEMRETTLLTSTQINTEBERWEGYGVQOVQVTSIGAAGAAAS 397
QY 10 ----- 9
Db 398 VSASASYAAEAATGAKVQDADKSAAVATVVAADVMAVRVVISAVQTAQRTTTTA 457
QY 10 ----- 9
Db 458 VHIQPAQEVKKEAKTAVTKVVVAADKAKEQELKSRKEIITTKQBQMHVTHEQIRKET 517
QY 10 ----- 9
Db 518 EKTIVPKVLSAAKAKEQETRISEIITKKQOVTOEALMKETRTVVVPKVIVATPKVKEQ 577
QY 10 ----- 9
Db 578 DLVSRGREGITTKREQVOITQEKMRKEAKETALSTIAVATAKAKEQETILTRTETMATRQ 637
QY 10 ----- 9
Db 638 EQIQVTHGKVDVGKKAABAVATVVAADVQARVREPREGHLEESYAQOTTLEYGYKERISA 697
QY 10 ----- 9
Db 698 AKVAEPPQRPASBPHVVPKAVPRVIOAPSETHIKITTDQGMHISQIKKTTDLTTERLV 757
QY 10 ----- 9
Db 758 HVDKRPRTASPHFTVSKISVPKTEHGYEASIAATLQKLSATSSAQKITKSVKAPT 817
QY 10 ----- 9
Db 818 VKPSETRVRAEPTLPQFPFADPTDYKSEAGVEVKEVGVSTGTTVREERFEVLHGHE 877
QY 10 -----CPAP-----ELGGSPVFL-----PP----- 25
Db 878 AKVTETARVPAPVEIPVPTPLVSLGNVTVIEGESVTLCHISGYPSPVTWYREDYQI 937
QY 26 ----- 25
Db 938 ESSIDFQITFQSGIARLMIREAFAEDSGRFTCSAVNEAGTVSTCYLAVQVSEFEKETT 997
QY 26 ----- 25
Db 998 AVTEKFTTEKRFVESRDVMTDSTLSEQAQGPGEPAAPYFITKPVVQKLVEGGSVVFQC 1057
QY 26 -----PKP----- 28
Db 1058 QVGNPKPHVYWKSGVPLTTGYRYKVSYNKQTGECKLVISMFTADDAGEYTIIVRNKHG 1117
QY 29 ----- 28
Db 1118 ETSASALLEADYELIMKSQEMLYQTQVTAQVQPEVGETAPGVEYSEYEKEQA 1177
QY 29 -----KDTLM----- 33
Db 1178 LIRKKMAKDTVVVTVTVEDQEFHISSEERLIKEIYRIKTTLEELLEEDGEEKMAVDI 1237
QY 34 ----- 33
Db 1238 SESEAVESGFDLRIKNRILEGNTVTHCKMCGYPLPKIANYKDGKRIKHGERYQMDFLQ 1297
QY 34 ----- 33

Db 1298 DGRASLRIPVVLPEDEGIYTAFAASNIKGNAICSGKLYVEAPALGAPTYIPTLEPVSRI 1357
QY 34 -----ISRTP----- 38
Db 1358 SLSPRSYSRSPIRMSPARMSPARMSPARMSPARMSPARMSPARMSPARMSPARMSPAR 1417
QY 39 ----- 38
Db 1418 VSPKCLEGANCDFDLKVVGPRMPETFWFDGQQLVNDYTHKVVIKEDGTOSLIIVPATPS 1477
QY 39 ----- 38
Db 1478 DSGEWTVAQNAGRSSISVILTVEAVEHQVPMFEVKLNKVNKIKESRLEMKVATGNP 1537
QY 39 ----- 38
Db 1538 NPDIVMLKNSDIIIVPHKYPKIRIEGTGKAALKIDSTVSQDSAWYTATATINKAGRDTRC 1597
QY 39 ----- 38
Db 1598 KVNVEVEFAPEPERKLIIPRGTYRAKEIAAPELEPLHLRYGQEQWEEGDLYDKKQOKP 1657
QY 39 ----- 38
Db 1658 FFKKLTSLRKRFGPAHFECRLTPISDPTMVVWHLHDGKPLEAANRLRMINFEGYCSLD 1717
QY 39 -----EUTC----- 42
Db 1718 YGVAYSRDSGIITCRATNKYGTDHTSATLIVKDEKSILVESQLPEGRKGLQRIEELERMA 1777
QY 43 ----- 42
Db 1778 HEGALTVTTDQKEKOKPDIVLYPEVPVLEGETARFCRVCTGYPOPKVWYLNGLQIRK 1837
QY 43 ----- 42
Db 1838 SKRFRVRDGIHYLDIVDCKSYDTGEVKVTAENPEGVIEHKVLEIQOQREDRSVLRRAP 1897
QY 43 ----- 42
Db 1898 EPRPEFHVHPGKLQFEVQKQVDRPVTETKEVVVKLRAERITHEKVPESEELRSKFKR 1957
QY 43 ----- 42
Db 1958 RTEEGYEAITAVELKSRKDESYEBELLRTKDELLHWTKELTEEEKKALAEKGKITIPT 2017
QY 43 ----- 42
Db 2018 FKPKIELSPSMEAPKIFERIQSQTVGQSDAHFRVVRVVGKPDCEWYKNGVKIERSDR 2077
QY 43 -----VVYDVSHED----- 51
Db 2078 IYWPEDNVNCELVIIRDVTAEDSASIMWKAINIAGETSSHAFLVQAKOLITFTQELQDV 2137
QY 52 -----PEVKFNWYVDGVEVH----- 66
Db 2138 VAKEKDTMATFECETSEPPFVKYKWKYKDGMEVHEGDKYRMHSDRKVFHLSLTITDTSDAED 2197
QY 67 -----NAKT----- 70
Db 2198 YSCVLVEDENVKTTAKLIVEGAVVEFVKELQDIEVPESYSGELECIVSPENIEGKWYHND 2257
QY 71 -----KPR----- 73
Db 2258 VELKSNKYTIITRRGRQNLTVKDVTKEDQGEYSFVIDGKTKTKLUMKPRPRTAILQGLS 2317
QY 74 ----- 73
Db 2318 DQKVCBGDIVQLEVKVSLESVEGVWMDGQEVQPSDRVHIVIDKQSHMLLIEDMTKEDAG 2377
QY 74 ----- 73

QY 202 ----- 201
Db 31358 PSQGELEILSISKSVTLQWEKPCDGGKEILGYWVEYRQSGDSAWKSNKERIKDQF 31417
QY 202 ----- 201
Db 31418 TIGGLEATEYEFVFAENETGLSPRRRTAMSIKTLTSGEAPGIRKEMKDVTTKLGEAA 31477
QY 202 ----- 201
Db 31478 QLSQOICRPLPDIKWYRFGKELIOSRYKMWSSDGRTHLTVMTEEQDEGVYTCIATNE 31537
QY 202 ----- 201
Db 31538 VGEVETSSKLLLOATPQHPGYPLKEKYGAVGSTLRLHVMYIGRPVPAMTWFHCKLILQ 31597
QY 202 ----- 204
Db 31598 NSENITIENTEYTHLVMKNVQRKTHAGKYKVQLSNVFGTVDAILDVEIQDKPKPTGPI 31657
QY 205 ----- 204
Db 31658 VIEALLKNSAVISWKPADDGGSWITNVYVEKCEAEGAELVSSAISVTTTCRIVNLTE 31717
QY 205 ----- 208
Db 31718 NAGYFRVSAQNTFGISDPLEVSSVILKSPFEKPGAPGKPTITAVTKDSCVVAWKPPAS 31777
QY 209 ----- 208
Db 31778 DGGAKIRNYLEKREKKONKWISVTTEIRETVFSVKNLIBGLEYEFRVKCNLGGSEW 31837
QY 209 ----- 208
Db 31838 SEISEPITPKSDVPIQAPHFKEELNLRVYQSNATLVCKVTGHPKPIVKWYRQKEIIA 31897
QY 209 ----- 208
Db 31898 DGLYRIOEFKGYHQLIIASVTDATVYQVRATNOGGSVSGTASLEVEVPKIHLPKT 31957
QY 209 ----- 208
Db 31958 LEGMAVHALRGEVSVKIPSPGKDPVITWQKQDLIDNNGHYQVIVTRSTSLSVFPNG 32017
QY 209 ----- 208
Db 32018 VERKDAGYVVCANRFGIDQKTVELDVADVPDPGRGVKVDSDVSNLVTWTEPASDGG 32077
QY 209 ----- 208
Db 32078 SKITNIVIEKATTAERWLRVGOARETRYTVINLFGKTSYOFVIAENKFGLSKPSBPSE 32137
QY 209 ----- 208
Db 32138 PTITKDKTRAMNDEEVDETVESMTKASHSSTKELYEKYMAIEDLGRGEFGIVHRCVE 32197
QY 209 ----- 212
Db 32198 TSSKKTYMAKVKVGTDOVLVKKELISILNIAHRNLIHLHESFESMEELVMIFETISGL 32257
QY 213 ----- 216
Db 32258 DIFERINTSAFELNREIVSVHOVCEALQFLSHNIGHFDIRPENIIYQTRRSSTIKII 32317
QY 217 ----- 216
Db 32318 EFGQAROLKPDNFRLLFTAPEYAPVHOVDVVSTATDMWSLGTLYVLLSGINPFLAE 32377
QY 217 ----- 218
Db 32378 TNOQIIENIMNAEYTFDEAFKEISIEAMDFVDRLLVKERKSRMTASEALQHPWLKOKIE 32437
QY 219 ----- 225
-----QKSLSL-----

Db 32438 RVSTKVIRTLKRRYYHTLIKDLNMVSAARISCGGAIRSQKGVSAKVKVASIEIGPV 32497
QY 226 ----- 225
Db 32498 SQQIMHAVGEEGHVYCKIENYDQSTQVTWYFGVROLENSEKEYEITYEDGVAILYVKD 32557
QY 226 ----- 225
Db 32558 ITKLLDDGTYRCKVNDYGEDSSYAEFLVKGREVVYDYCRRTMKKIKRRTDMRLLRPP 32617
QY 226 ----- 225
Db 32618 EFTLPLYNKATVAVGVNVRFGVTITVHPEPHVTWYKSGOKIKPGDNDKXYTPESDKGLQL 32677
QY 226 ----- 225
Db 32678 TINSVTTDDAEYTVARNKYGEDSKAKLTVTLHPPPTDSTLRPMFKRLLANAECQEQ 32737
QY 226 ----- 225
Db 32738 SVCFEIRVSGIPPPTLKWKGQPLSLGPNIEIIHEGLDYYALHTRDLPEDTGYRYRTA 32797
QY 226 ----- 225
Db 32798 TTAGSTSCQAHLOVERLYKKQEFKKEEHERHVQKQIDKTLRMAEILSGTESVPLTQV 32857
QY 226 ----- 225
Db 32858 AKEALREAAVLYKPAVSTKTVKGEFRLEIEBKKEERLMPYDVPEPRKYKQTTIEEDQR 32917
QY 226 ----- 228
Db 32918 IKQFVPMKMKYKIRDOYEMPGKLRVQKRPKIRLSRWEQFYVMPLEPRITDQYRPK 32977
QY 229 ----- 228
Db 32978 WRIPKLSQDDLEIVRPARRTTSPDYDFYRPRRRSLGDISDEELLIPIDDOYLAMKRTEE 33037
QY 229 ----- 228
Db 33038 ERLLEELELGFASPPSRSPHFLSSLRYSSPQAHVKVETRKDFRYSTYHIPTKAE 33097
QY 229 ----- 228
Db 33098 ASTSYAELRERHAQAAVROPKQORIMAREDEELLRPVTTTQHLSEYKSELDPMKSEEK 33157
QY 229 ----- 228
Db 33158 SRKKSRRQREVTEITEIEEYEISKHAQRESSSASRLRRRRSLSPITYIELMRPVSELI 33217
QY 229 ----- 228
Db 33218 RSRPQAEYEDDTERRSPTPTERTPRSPSPVSSERSLSPERSAREDFISRYESMKAAL 33277
QY 229 ----- 228
Db 33278 KTQKTSERKEYVLSQOPFTLDHAPRITLMRSHRVPQGNTRFTILNVOSKPTAEVKWYHN 33337
QY 229 ----- 228
Db 33338 GVELOESSKIHTNTSGVLTJLEILDCHTDSDGTYRAVCTNYKGEASDYATLDVTGGDYTT 33397
QY 229 ----- 228
Db 33398 YASQRDEEVPVFPPELTRTEAYAVSSFKKTSEMEASSSVREVYKQMTETRESLSYEH 33457
QY 229 ----- 228
Db 33458 SASAEMKSAALEKSLSEKSTTRKIKYTLAARILTKPRSMTVYEGESARFSCDTDGEPVP 33517
QY 229 --DWL----- 231

Db 29138 RFKVTGLTEGNEYEFHVMAENAGVGPASGISRLIKREPVPNPGPTWVKVTDTSKTTV 29197
QY 183 ----- 182
Db 29198 SLEWSKPVFDGMEIIGYIIEMCKRADLGDWHKVNACVKTRYTVTDLOAGEEYKFRVSA 29257
QY 183 ----- 182
Db 29258 INGAGKGDCEVTGTIKAVDRLETAPELDIDANFKQTHVVRAGASIRLFIAOGRPTTAV 29317
QY 183 ----- 182
Db 29318 WSKPDSNLSLRADITHDTSFSTLAVENCNRNDAGKYTLTVENNSGSKSITFTVKVLDPG 29377
QY 183 ----- 182
Db 29378 PPGPTEKDVTRGSATLMDAPLLDGGARHHYVVEKREASRWQVISEKTRQIFKVN 29437
QY 183 ----- 182
Db 29438 DLAEVGPYIFRVSANVEYGVGEPEYEMPEPIVATEQPAPPRRLDVVDTSKSSAVLAWLKP 29497
QY 183 ----- 182
Db 29498 HDGSRITGYLLEMRQKGSDFWEAGHTKOLTFTEVERLVEKTEYEFVRKAKNDAGYSEPR 29557
QY 183 ----- 182
Db 29558 EAFSSVILKEPQIEPTADLTGINTLITCKAGSFTTDVPISGRPAPKVTKWLEMLKE 29617
QY 183 -----GSFFL----- 187
Db 29618 TDRVSITTKDRTTLTVKDSMRGDSGRYFLTLENTAGTKFTSVVVVIGRPGVTPGIEV 29677
QY 188 ----- 187
Db 29678 SSVSAESCVLSWGEKPDGSGTEITNYIVEKRESGTTAWOLVNSVKRTQIKVTHLTKYME 29737
QY 188 ----- 187
Db 29738 YSRVSENRFVSKPLESAPIIAEHFPVPPSAPTRPEVYHVSAWNSIRWEEPHYDGG 29797
QY 188 ----- 187
Db 29798 KIIGYWEKKERNLILWKENKVPCLCNKYVTGLVELEYQFTYALNAGVSKASEAS 29857
QY 188 ----- 187
Db 29858 RPIAQNPDVAPGRPEVTDVTRSTVLSWAPAYDGGKVVGYIIEKKPVSEVGDGRWLK 29917
QY 188 ----- 187
Db 29918 CNVTIYSDNFFVTALSEGDTEYFRVLAKNAGVISKSESTGPVTCRDEYAPPKAELDA 29977
QY 188 ----- 187
Db 29978 RLHGLVTRAGSDLVDAAGVKPEPKIITWKGDELCEKVSLOYTGKRATAVIKFC 30037
QY 188 ----- 187
Db 30038 DRSDSGKYTLTVKNASGTRKAVSVMKVLDSPGPGKLTVSRTVQEKCTLAWSLPQEDGGA 30097
QY 188 ----- 187
Db 30098 EITHYIVERRETSRLNWIVEGECPTLSYVVTRLIKNNEYIFRRAVKNYGFVPESEP 30157
QY 188 ----- 187
Db 30158 IVARNSTIPSPGPIPEEVTGKEHIIIIQWTKPESDGGNEISNYLVDKREKSLRWTRVN 30217
QY 188 ----- 187
Db 30218 KDYVVYDTRKVTSLMEGCDYQFRTAVNAAGNSBPSEASNFISCREPSYTPGPPSAPRV 30277

QY 188 ----- 187
Db 30278 VDTTKHSISLAWTKPMYDGGTDIVGYVLEMOEKDQWVRVHTNATIRNTEFTVPLKMG 30337
QY 188 ----- 187
Db 30338 QKYSFRVAANVWKGMSSESESAIEIEPVERIEIPDELAADDLKKTVTTIRAGASLRMLSV 30397
QY 188 -----YSKLTVDK----- 195
Db 30398 SGRPPPVITWSKQIGIDLASRAIDTTESYLLIVDKVKNRYDAGKYTIEAENOSGKKSATV 30457
QY 196 ----- 195
Db 30458 LVKYDTPGPCPSVKKEVSRDSVTITWEIPTIDGAPVNNYIVEKREAAAMRAFKTVTK 30517
QY 196 ----- 195
Db 30518 CSKTLYRISGLVEGTMYYFRVLPENIYGIGEPCTSDAIVLSEVPLVPAKLEVVVDVKST 30577
QY 196 ----- 195
Db 30578 VTLAWEKPLYDGGSRSLGVLVLEACKAGTERMMKVVTLLKPTVLEHTVTSLSNEGEQYLFIR 30637
QY 196 ----- 195
Db 30638 AQNEKGVSEPRETVTAVTQDLRVLPTIDLSTMPQKTHVPAGRPVELVPIAGRPPPA 30697
QY 196 ----- 195
Db 30698 SWFFAGSKLRESERVTVETHTKVAKLTIRETTIRDTGEYTLKKNVTGTTSETIKVILD 30757
QY 196 ----- 195
Db 30758 KPGPTGPIKIDEIDATSIISWEPPELDGAPLSGVVVFQORDAHRPCWLPVSESVRST 30817
QY 196 ----- 195
Db 30818 FKFTRLTEGNEYFRVAATNRFVIGSYLQSEVIECRSSIRIPGPETLQIFDVSROGRTL 30877
QY 196 ----- 195
Db 30878 TWYPEDDGGSQVTGYIVERKEVRADRWVRNVKVPVTMTRYSTGLTEGLEYEHRVTAIN 30937
QY 196 ----- 195
Db 30938 ARGSGKPSRSPKPIVAMDPIAPPGKQONPRVTDTRTSVSLAWSVPEDEGSKVTGYLIE 30997
QY 196 ----- 195
Db 30998 MOKVDOHEWTKCNTTPTKIREYTLTHLPGAERYFRVLACNAGGPGPAEVPGTVKVTEM 31057
QY 196 -----SRWQOG----- 201
Db 31058 LEYDPYELDERYQEGIFVRQGGVIRLTIPIKKGKPPICKWKKEGODISKRAMIATSETHT 31117
QY 202 ----- 201
Db 31118 ELVIKEADRGDSCTYDLVLNCKCGKKAIVKVRIGSPNSPEGPLEYDDIQVRSVRVSWR 31177
QY 202 ----- 201
Db 31178 PPADGGADILGYLERREVPKAAWYTIIDSRVGTSLVVKGLKENVEYHFRVSAENQFGI 31237
QY 202 ----- 201
Db 31238 SKPLKSEEPVTPKTLPLNPPEPPSNPPELVDVTKSSVLSWSRPKDDGSRVTGYIERKE 31297
QY 202 ----- 201
Db 31298 TSTDKVRHINKTOITTTMTVTGLVDPDAEYQFRIIAQNDVGLSETSPASEPVVCKDPFDK 31357

QY 176 ----- 175
Db 26978 CEYEVVAENAGLSLPSETSPILRAEDPVFLPSPKPKIVDSGKTTITIAWKPLFD 27037
QY 176 ----- 175
Db 27038 GGAPITGYTVEYKSDDTDMKTSIQSLRGTEYTSIGLTTGAEYVFRVSKVKNVKGASDPDS 27097
QY 176 ----- 175
Db 27098 SSDPOIAKEREEPLFDIDSEMRKTLIVKAGASFTMTVPFRGRPVPNVLMWSPDPTDLRTR 27157
QY 176 ----- 175
Db 27158 AYVDTTDSRTSLTIENANRNDGKYTLTIQNVLSAASLTLVVKVLDLTPGPTNITVQDVT 27217
QY 176 ----- 175
Db 27218 RESAVLSWDVPENDGGAPVKNYHIEKREASKAMVSVTNNCNRLSKYKVTNLQEGAIYYFR 27277
QY 176 ----- 175
Db 27278 VSGNEFGVIPAETKEGVKITEKPSPEKLGVTISKDSVSLTWLKPEDGGSRIHVY 27337
QY 176 ----- 175
Db 27338 VEALEKGOKNWKCAVAKSTHHVYSGLRENSYFPRVFAENQAGLSDPRELLLPVLIKEQ 27397
QY 176 ----- 175
Db 27398 LEPPEIDMKNPSSHVVYVRAGSNLKVDPISGKPLPKVTLSDRGVPLKATMRNFTEITAE 27457
QY 176 ----- 175
Db 27458 NLTNLKESTADAGRYEITAANSSGTTKAFINIVLDRPGPTGPVVISDITEESVTLK 27517
QY 176 ----- 175
Db 27518 WEPPYDGSQVTNVIILLKRETSATVTEVSATVARTMKYMKLTTGEEYQFRIKAENRF 27577
QY 176 ----- 175
Db 27578 GISDHDSACVTVKLPYITPGPPSTPWVNTVRESITVWHEPVSNAGSVAVGYHLEMD 27637
QY 176 ----- 175
Db 27638 RNSILWQANKLVIRTHFKVTTISAGLIYEFVVAENAGVGKPSHPSPVLADACEP 27697
QY 176 ----- 175
Db 27698 PRNVRTDISKNSVLSWQQPAFDGSKIITGYIIVERDLDPGRWTKASFTNVTETQFIIS 27757
QY 176 ----- 175
Db 27758 GLTONSQYEFVRFARNAGVSGISNPSEVGPITCIDSYGGPIDLPLEYEVVKYRAGTSV 27817
QY 176 ----- 175
Db 27818 KLRAGISGPAPTIEWYKDDKELOTNALVCVENTDLSILIKDADRNSGCEYELKURNA 27877
QY 176 ----- 175
Db 27878 MGSATIRVOILDKPGPPGGPIEFKVTAEKITLLWRPPADGAKITHYIVEKRETSR 27937
QY 176 ----- 175
Db 27938 VVMSWSEHLEECIITTTKIIKNEYIFRVRANKYIGEPLESDSVVAKNAFVTPGPPG 27997
QY 176 ----- 175
Db 27998 IPEVTKITKNSMTVVWSRPIADGSDISGYFLEKRDKSLGWFKVLKETIRDTROKVTGL 28057
QY 176 ----- 175

Db 28058 TENSQYQYRVCAVNAAGOPFSEPFYKAADPIDPPGPAKIRIADSTKSSITLWGSKP 28117
QY 176 ----- 175
Db 28118 VYDGSVAVTGYVVEIROGEEEEETVSTKGEVRTTEYVVSNLKPGVNYFRVSAVNCAGQ 28177
QY 176 ----- 175
Db 28178 GEPIEMNEPVQAKDILEAPEIDLDVALRTSVIAKAGEDVQVLI PFKGRPPPTVTRKDEK 28237
QY 176 ----- 175
Db 28238 NLGDARYSIENTDSSLLTIPQVTRNDTGKYLITENGVGEPKSTSVSVKVLDTPAACQ 28297
QY 176 ----- 175
Db 28298 KLQVHVSRGVTLLWDPLIDGGSPIINYVIEKRDATKRTWSVVSHKCSSTSFKLIDL 28357
QY 176 ----- 175
Db 28358 EKTPTFFRVLAENEIGIGEPCEETPEPVKAAEVPAPIRDLMSKMDSTKTSVLSWTKPDFG 28417
QY 176 ----- 175
Db 28418 GSVITEYVVERKGEQWTHAGISKTCEIEVSQLEQSVLEFRVFAKNEKGLSDPVTIG 28477
QY 176 ----- 175
Db 28478 PITVKELIITPEVDLSDIPGAQVTVRIGHNVHLELPYKGPSPISWLKDLGLPKESEFV 28537
QY 176 ----- 175
Db 28538 RFSKTENKITLSIKNAKKEHGKYTVILDNAVCRIVAPITVITLGPSPKPGPIRDEIK 28597
QY 176 ----- 175
Db 28598 ADSVILSWDVPEDNGGGEITCYSIEKRETSQTNMKWCVSSVARTTKVPNLVKDAEYQFR 28657
QY 176 ----- 175
Db 28658 VRAENRYGVSQPLVSSIIIVAKHQFRIPGPGKPVIIYNTSDGMSLTWDAPVYDGGSEVTG 28717
QY 176 ----- 175
Db 28718 FHVEKKERNILWOKVNTSPISGREYRATGLVEGLDYQFRVYAENSAGLSPSPDSKFTL 28777
QY 176 ----- 175
Db 28778 AVSPVDPPTDYIDVTRETITLKNPPLRDGGSKIYGVYIEKROGNERWVRCNFTDYSE 28837
QY 176 ----- 178
Db 28838 QYTVTGLSPGDRYEFRIIARNVAGTISPPSQSGIIMTRDENVPPIVEFGPEYFDGLII 28897
QY 179 ----- 178
Db 28898 KSGESLRIKALVQGRPVPRVTWFKDGVIEKRMNMWETDVLGSTSLFVRDATRDRHGVYT 28957
QY 179 ----- 178
Db 28958 VEAKNASGAKAEIKVKVQDTPGKVVGPIRFTNITGEKMTLWMDAPLNDGCAPITHYIIE 29017
QY 179 ----- 182
Db 29018 KRETSRLAWALIEDKCEAQSYTAIKLINGNEYQFRVSAVNFVGRPLDSDPWWAQIOYT 29077
QY 183 ----- 182
Db 29078 VPDAPGIPESNITGNSITLTWARPESDGGSEIOQYILERREKKSTRWVKVSKRPISET 29137
QY 183 ----- 182

Db 24758 AKDSMTVCNRPDSGGSEIIIGYIIVEKDRSGIRWIKCNKRITDLRLRVTLGTEDHEYE 24817
QY 168 ----- 167
Db 24818 FRVSAENAGVGPSPATYYKACDPVKPGPTNAHIVDTTKNSITLAWGKPIYDGGSE 24877
QY 168 ----- 167
Db 24878 ILGYVVEICKADEEWEQIVTQPTGLRVTRFEISKLEHQEKIRVCALNKVGLGEATSPV 24937
QY 168 ----- 167
Db 24938 GTVKPEDKLEAPELDSELRKGIIVRAGGSARIHIPKGRPTPEITWSREGEFTDKVO 24997
QY 168 -----PEN----- 170
Db 24998 IEKGVNYTOLSIDNCDNRNDAGRYILKLENSGSKSAFVTVKVLDTGPPONLAVKEVRKD 25057
QY 171 ----- 170
Db 25058 SAFLVNEPPIIDGGAUKVNYVIDKRESTRKAYANVSSKSTSKFVENITEGAIYFRVM 25117
QY 171 ----- 170
Db 25118 AENERGVGPVETVDAVKAAPPPPGKVTLDVSQTSASLWKEPDEHDGGSRLGYVVE 25177
QY 171 ----- 170
Db 25178 MQPKGTEKWSIVAESKVCNAVVTGLSSGOEYQFRVKAYNEKSKSDPRVLGVPVIAKDLTI 25237
QY 171 ----- 170
Db 25238 QPSLKLPFNTYSIQAGEDLKIEIPVIGRPRPNISWVKDGEPLKQTRVNVEETATSTVLH 25297
QY 171 ----- 170
Db 25298 IKEGNKDDFGKYTVATNSAGTATENLSVIVLEKPGPPVGVPRFDEVSADFVVISWEPPA 25357
QY 171 ----- 170
Db 25358 YTGCGQISNYIVEKRDTTTTTHWMSATVARTTIKITLKTGTEYQFRIFAENRYGKSAP 25417
QY 171 ----- 170
Db 25418 LSKAVIYOYPKEPGPGTPEVTISIKDOMLVQWHEPVNDGGTKIIGYHLEQKEKNSIL 25477
QY 171 ----- 170
Db 25478 WYKLNKTPIDTKFKTTGLDEGLEEYFKVSAENIVGIGKPSKVSCEFVARDCDPPGRPE 25537
QY 171 ----- 170
Db 25538 AIVITRNNVTWKWKPAYDGGSKITGYIVEKKDLDPGRMKASFNVLETFVTSGLVED 25597
QY 171 ----- 170
Db 25598 QRYEFRVIARNAAGNFSEPSDSSGAITARDEIDAPNASLDPKYKDVIVVHAGETVLEAD 25657
QY 171 ----- 170
Db 25658 IRGKPIPDVWWSKSGKELEETAARMEIKSTIQKTTLVVKDCIRTDGGQYILKLSNVGGTK 25717
QY 171 ----- 170
Db 25718 SIPITVKVLDPRCPPEGPLKVTGVTAEKCYLAWNPPLDGGANISHYIIEKRETSRLSWT 25777
QY 171 -----NYKT----- 175
Db 25778 QVSTEVOALNYKVTLLPGNEYIFRVMVAVNYKYGIGEPLESGPVTACNPKPPGPPSTPEV 25837
QY 176 ----- 175
Db 25838 SAITKDSWVVTWARPVDGGEIEGYILEKKRKEGVWTKCNKKTLDLRLRVTLGTGEH 25897

QY 176 ----- 175
Db 25898 SYEFRVAENAGVGPSEPSVYFACDALYPGPPSNPKVTDTSRSSVSLAWSKPIYDG 25957
QY 176 ----- 175
Db 25958 GAPVKGYYVVEKAAADEWTTCTPPTGLOGKQFTVTCLKENTEYFRICAINSEGVEPA 26017
QY 176 ----- 175
Db 26018 TLPGSVVAQERIEPPEIELDADLRKVVLURASATLRLFTIKGRPEPEVKEKAEGILTD 26077
QY 176 ----- 175
Db 26078 RAQIEVTSFTMLVIDNVTRFDSGRYNLTLENNSGSKTAFVNVVLDSPSPVNLTIREV 26137
QY 176 ----- 175
Db 26138 KKDSVTLSEPPPLIDGGAKITNYIVEKRETTTKAYATITNNCTKTTFRIENLOEGSYIF 26197
QY 176 ----- 175
Db 26198 RVLASNEYCIGLPAETTEPVKYSEPLPGRVTLVDVTRNTATIKWEKPESDGGSKITGY 26257
QY 176 ----- 175
Db 26258 VVEMQTKGSEKWTCTQVKTLEATISGLTAGEEYVFRVAAVNEKGRSDPRQLGVPVIARD 26317
QY 176 ----- 175
Db 26318 IEIKPSVELPFTFNVKAREQLKIDVPFGRPOATVNMWKKDQTLKETTRVNVSSKVTY 26377
QY 176 ----- 175
Db 26378 SLSIKEASKEDVGTVELCVSNSAGSITVPTIIVILDRPPOGPIRIDEVSCDSITISWNP 26437
QY 176 ----- 175
Db 26438 PEYDGGQISNYIVEKKTETSTTHWIVSOAVARTSIKIVRLTTGSEYQFRVCAENRYGKS 26497
QY 176 ----- 175
Db 26498 SYSSESAVVAEYFPSPGPPGTPKVVHATKMTLVMTQVVPVNDGGSRVIGYHLEYKERS 26557
QY 176 ----- 175
Db 26558 ILWSKANKTLIADTQMKVSGLDEGLMEYRVAENIAGIGCKSCPEVPARDPCDPPGQ 26617
QY 176 ----- 175
Db 26618 PEVTNITRKSLSKWSKPHYDGGAKITGYIVERRELDPGRWLKCNNTNIQETYFEVTELT 26677
QY 176 ----- 175
Db 26678 EDQRYEFRVFAARNAADSVSESTGPIIVKDDVEPRVMDVKFRDVIIVVRAGEVLKIN 26737
QY 176 ----- 175
Db 26738 ADIAGRPLPVISWAKDGEIEERARTEIISTDNHTLLTVKDCIRRDYGOYVLTLLKNVAGT 26797
QY 176 ----- 175
Db 26798 RSVAVNCKVLDKPPGPPAGPLEINGLTAECSLSGWRQPDGGADIDYIYVEKRETSRLAW 26857
QY 176 ----- 175
Db 26858 TICGELOMTSCKVTLLKNGEYIFRVTGVNRYGVEPLESVAIKALDPFTVPSPPTSLE 26917
QY 176 ----- 175
Db 26918 ITSVTKESMTLWNSRPESDGGSEISGYIIERREKNSLRVRVNNKPVYDLRVKSTGLRGG 26977

QY	168	-----	167
Db	22598	MVVCWGHDPDGGSEILINIVERRDKAGQRIWKCNKKTLDLRYKVSGLTGEGHEYFRIM	22657
QY	168	-----	167
Db	22658	AENAAGISAPSPSPFFYKACDVFVKPGPNRPVLDTSRSSISIAMNKPIYDGGSEITGY	22717
QY	168	-----	167
Db	22718	MVEIALPEDEWQIVTPPAGLKATSYITGLTENQEKIRIYAMNSEGLGEPALVPGTPK	22777
QY	168	-----	167
Db	22778	AEDRMLPEIELDADLRKVVITIRACCTLRFLFVPIKGRPAPEVKWARDHGESLDKASIEST	22837
QY	168	-----	167
Db	22838	SSYTLIVGNVRPDSGKIILTVENSSGSKSAFVNRVLDTPGPPDLKVKVETKTSVTL	22897
QY	168	-----	167
Db	22898	TWDPPLLDGGSKIKNYIVEKRESTRKAYSTVATNCHTSMKWDLQLEGCSYIFRVLAENE	22957
QY	168	-----	167
Db	22958	YGIGLPAETAESVKASERPLPGKILMDVTRNSVLSWEKPEHDGGSRLGIYIVEMQTK	23017
QY	168	-----	167
Db	23018	GSDKWATCATVKVTEATITGLIOGEYSFRVSAQNEKGISDPRQLSVPIAKDLVIPPAF	23077
QY	168	-----	167
Db	23078	KLLFNTFTVLAGEDLKVDVDFIGRPTPAVTWHKDNVPLKOTTRVNAESTENNSLLTIKDA	23137
QY	168	-----	167
Db	23138	CREVDGVHVVKLTNSAGEAIELTNVILDKPGPTGPVKMDEVTAADSIILSWGPPKYDGG	23197
QY	168	-----	167
Db	23198	SSINNYIVEKRDSTTTWQIVSATVARTTIKACRLKTGCEYQFRIAAENRYGKSTYLNSE	23257
QY	168	-----	167
Db	23258	PTVAQYPFKVPGPPTPVVTLSSRDSMEVQWNEPISDGGSRVTGYHLERKERNISILWVKL	23317
QY	168	-----	167
Db	23318	NKTIPIQTKFTTGLEBGEVEFRVSAENIVGIGKPSKVSCEYVARDPCDPPGRPEALIV	23377
QY	168	-----	167
Db	23378	TRNSVTLOWKKPYDGGSKITGYIVEKKELPEGRWMKASFTNIIDTHFEVTGLVEDHRYE	23437
QY	168	-----	167
Db	23438	FRVIARNAAGVFSEPESTGAITARDEVDPPIRISMDPKYKDTIVVHAGESFKVDADIYK	23497
QY	168	-----	167
Db	23498	PIPTIOWIKDQELSNLARLEIKSTDFATSLSVKDAVRVDSGNVILKAKNVAGERSVTVN	23557
QY	168	-----	167
Db	23558	KVLDPRGPPGPGWISGVTAERKCTLAWKPPLODGGDIINIYIVERRETSRLVMTVV DAN	23617
QY	168	-----	167
Db	23618	VQTLSCVKTKLEBNEYTFRIMAVNKYGVGEPLSEBPVAKNPFVVPDAPKAPVETVTVK	23677
QY	168	-----	167

Db	23678	DSMIVVWERPASDGGSEILGVLEKRDKEGIRWTRCHKRILGELRLRVTGLIENHDEYFR	23737
QY	168	-----	167
Db	23738	VSAENAAGLSEPSPPSAYQKACDPIYKPGPNPNPKVIDITRSSVFLSWSKPIYDGGCEIQ	23797
QY	168	-----	167
Db	23798	GYIVEKCDVSVGWTMCTPPTGINKTNIEVEKLLLEKHEYNFRICAINKAGVGEHADVPGP	23857
QY	168	-----	167
Db	23858	IIVEEKLEAPDIDLLELRKIIINIRAGGSLRFLFVPKGRPTPEVKMGKVDGEIRDAAIID	23917
QY	168	-----	167
Db	23918	VTSTFTSLVDNVRNRYDSGKYITLTLENSSGTKSAFVTVRVLDTPSPVNLKVTEITKDSV	23977
QY	168	-----	167
Db	23978	SITWEPPLDGGSKIKNYIVEKREATRKSAAVVTNCHKNKSWKIDLOEGCSYIFRVTAE	24037
QY	168	-----	167
Db	24038	NEYGIGLPAQTADPIKVAEVPQPPGKITVDVTRNSVLSWTKPEHDGSGKIIQYIVEMQ	24097
QY	168	-----	167
Db	24098	AKHSEKSECARVKSQAIVITNLQTGEEYLFRRVVAVNEKGRSDPRSLAVPIVAKDLVIEP	24157
QY	168	-----	167
Db	24158	DVKAPAFSSYVQVQDULKIEVPISSGRPKPITWTKQGLPLKQTRINVTDSLOTLTUSIK	24217
QY	168	-----	167
Db	24218	ETHKDDGGQYITVANVVVGQKTASIEIVTLDKPDPKPGVKFDDVSAESITLSWNPPLYT	24277
QY	168	-----	167
Db	24278	GGQOITNYIVQKRDTTWTWVDSATVARTTLKVKLTGTGYQFRIFAENRYGQSPALE	24337
QY	168	-----	167
Db	24338	SDPIVAQYPIKEPGPGTPTATAIKSDSMVIQWHEPVNNGGSPVIGYHLERKERNISILWT	24397
QY	168	-----	167
Db	24398	KVNKTIIDHTQFKAQNLLEGIEYEFVYAENIVGVGKASKNSECYVARDPCDPPGTPEPI	24457
QY	168	-----	167
Db	24458	MVKRNEITLOWTKPVYDGGSMITGYIVEKRDLPDGRWMKASFTNVIETQFTVSGLTEDOR	24517
QY	168	-----	167
Db	24518	YEFRTAKNAAGAIKSPSSTGPIITAKDEVLPRIKSDPKFRDTIIVVNAGETFRLEADVH	24577
QY	168	-----	167
Db	24578	GRPLPTIEWLRGDKIEESARCEIKNTDFKALLIVKDAIRIDGGQYILLRASNVAGSKSFP	24637
QY	168	-----	167
Db	24638	VNVKVLDRPGPPGPGVQVTGVTSEKSLTWSPPPLQDGGSDISHYVVEKRETSRLAWTVVA	24697
QY	168	-----	167
Db	24698	SEVVTNSLKVTKLEBNEYVFRIMAVNKYGVGEPLSEAPVLMKNPFVLPGPSKLEVTNI	24757
QY	168	-----	167

Db 20378 KLIEGHEYQFRICAENKYGVGDVFTFTEPAIAKPNYDPPGRCDDPPVISNITKDHTMTVSWKP 20437
QY 154 ----- 153
Db 20438 PADDGSPITGYLLEKRETOAVNWKVNRPPIERTLKATGLOEGTEYEFRTAINKACP 20497
QY 154 ----- 153
Db 20498 GKPSDASKAAAYARDPOYPAPPAPFPKVDYDTRSSVSLSWGKPAYDGGSPILGYLVEVKRA 20557
QY 154 -----YPSDIAY- 160
Db 20558 DSDNNVRCNLPONLOKTRFEVTLGMDTQYQFRVAVNKNIGYSDPSDVPDKHYPKDILIP 20617
QY 161 -----|| || | 160
Db 20618 PEGELDADLRKTLILRAGVTWRLYVPVKGRPPPKITWSKPNVNLDRIGLDIKSDFTDFTE 20677
QY 161 ----- 160
Db 20678 LRCENVNKYDAGYILTLNENSCGKKEYTIVVKVLDTPGPPVNVTVKEISKDSAYVTWEP 20737
QY 161 ----- 160
Db 20738 IIDGSPILNYYQKRDAERKSNSTVTTECKSTSPRVANLEBEGKSYFFRVFAENYIGID 20797
QY 161 ----- 160
Db 20798 PGETRAVAKSOTPGPVVDLKVRSYSKSSCSIGWKKPHSDGSGRIIGYVVDFTLTKENKQ 20857
QY 161 ----- 160
Db 20858 RYMKSLQYSAKOLTEGKEYTFRVSAENENGEPTSEITVVARDVVDVAPDLKGLPDL 20917
QY 161 ----- 160
Db 20918 CYLAKENSFRLLKIPKGPAPSVWKKGEDPLATDTRVSVESAVNTLIVYDCQKSDA 20977
QY 161 ----- 160
Db 20978 GRYTITLKNVACTKEGTISIKVVGKPGIPTGPIKFEVTAEAMTLKWPAPKDDGSEITN 21037
QY 161 ----- 160
Db 21038 YILEKRDSVNNKWTCASAVQKTTFRVRLHEGMEYTFRVSAENKYGVGEGLKSEPIVAR 21097
QY 161 ----- 160
Db 21098 HPFDVPDAPPPNIVDVRHDSVSLTWTDPKKTGGSPITGYHLEFKERNLLWKRAKTP 21157
QY 161 ----- 160
Db 21158 RMRDFKVTGLTEGLEFEFRVMAINLAGVGPSPSEPVALDPIDPPGKPEVINITRNSV 21217
QY 161 ----- 160
Db 21218 TLIWTEPKYDGHKLTGYIVEKRDLPSKSMKANHVNVPECAFTVTDLVEGGKYEFIR 21277
QY 161 ----- 160
Db 21278 KNTAGNAPSSESTETIICKDEYEAPTIVLDPTIKDGLTIKAGDVIIVLNAISILKPLPK 21337
QY 161 ----- 160
Db 21338 SWSKAGDIRPSDITQITSTPTSSMLTIKYATRKDAGEYITITATNPFCTKVEHVKTVL 21397
QY 161 ----- 160
Db 21398 DVPGPPEVSEISVSAEKATLTWTPLEDGGSPIKSYILEKRETSRLLTWTVVSEDIQSCR 21457
QY 161 ----- 160
Db 21458 HVATKLIQNEYIFRVSANVHYKGEPVQSEPVKAVDRFGPPGPEKPEVSNVTKNTATV 21517

QY 161 ----- 160
Db 21518 SWKRPVDDGGSEITCYHVERREKKSILRWVYAIKTPVSDLRCKVTGLQEGSTYEFVRSAEN 21577
QY 161 ----- 160
Db 21578 RAGIGPSEASDVLMDAAAYPPGPSNPHVDTTKKSASLAWGRPHYDGGLEITGVYVE 21637
QY 161 ----- 160
Db 21638 HOKVGDEAWIKDTTCTALRITQFVVPDLQTKENYFRISAINDAGVGEPVAPIDPVEIVER 21697
QY 161 ----- 160
Db 21698 EMAPDFELDAELRRTLIVVRAGLSIRIFVPIKGRPAPEVTWTNDKINLKNRANIENTESPT 21757
QY 161 ----- 160
Db 21758 LLIIPENRYDTGKFVMTIENPAGKSGFVNVRVLDTPGPVNLRLRPTDITKDSVTLHMDL 21817
QY 161 ----- 160
Db 21818 PLIDGSRITNIVIVEKREATRKSYSYATTKCHKCTYKVTGLSEGCEYFFRVMAENEYIG 21877
QY 161 ----- 160
Db 21878 EPTETTEPVKASEAPSPDPSLNMIDITKSTVSLAWPKPKHDGSKITGYVIEAQRKGSQ 21937
QY 161 ----- 160
Db 21938 WTHITTVKLECVVRNLTEGEYTFQVMVNSAGRSAPRESRPVIVKEQTMPLPELDLRI 21997
QY 161 ----- 160
Db 21998 YQKLVIAKAGDNKVEIPVLGRPKPTVTKGQDQILKQTVYNFETTATSTILNINECYR 22057
QY 161 ----- 160
Db 22058 SDGSPYPLTARNIVGEGVDITIQVHDIPGPPTGPIKFDEVSSDFVTFSWDPPENDGGVP 22117
QY 161 ----- 160
Db 22118 ISNYVEMRQDSTTWVELATTVIIRTYKATRLTTGLEVOFRVKAQNRXGVGPGITSACI 22177
QY 161 ----- 160
Db 22178 VANYFPKVPGPPTQVTAVTKDSMTISWHHEPLSDGGSPILGYHYVERKERNGLMQTVSK 22237
QY 161 ----- 160
Db 22238 ALVPGNIFKSGSLTDGIAYEFVRVIAENMAGSKSPKSEPMALDPIPPGKPVPLNIR 22297
QY 161 ----- 160
Db 22298 HTVTLLKAKPEYTGFKITSYIVERKDLPNGRWLKANFSENFTEVTSGLTEDAAEYR 22357
QY 161 ----- 160
Db 22358 VIAKNAAGAI SPPSESDAITCRDDVEAPKIKVDVKFKDVTILKAGEAFRLEADYSGRPP 22417
QY 161 -----EWESNGO- 167
Db 22418 PTMEWSKDGKELEGTAKLEIKIADFSTNLVKNKSTRDSDGAYTLTATNPGFAKHIFNVK 22477
QY 168 ----- 167
Db 22478 VLDRGPPEGPLAVTEVTSEKCVLSWFPPLDDGGAKIDHYIVOKRETSRLATNVTNASEVQ 22537
QY 168 ----- 167
Db 22538 VTKLKVTKLLKGNXYIFRVMAYNKYGVGEPESEPVLA VNPYCPDPKPNPEVTITKDS 22597

QY	154	-----	153
Db	18218	VENPAGSKTVSVKVLVDKPGPRDLVSEIRKDSCLYTHKEPFLDDGGSVITNYVVERRD	18277
QY	154	-----	153
Db	18278	VASAQMSPLSATSKKKSHFAKHLNEGQYLFRAAENQYGRGPFVETPKPIKALDPLHPP	18337
QY	154	-----	153
Db	18338	GPPKDLHHVDVDTKTEVSLVWNNKPDROGGSPITGYLVEYQEGQDWIKFKVTNLCVYT	18397
QY	154	-----	153
Db	18398	GLQOGKTYRFRVKAENIVGLGLPDTTIPIECQEKLVPPSVELDVKLIEGLVVKAGTIVRF	18457
QY	154	-----	153
Db	18458	PAIIRGVVPVPTAKWTTDGSSEIKTDEHYTVETDNFSSVLTIKNCRRDTGEYQITVSNAAG	18517
QY	154	-----	153
Db	18518	SKTVAVHLTVLDVPGPPTGPDINILDTPEHMTISWQPPKDDGSPVINIYVEKQDTRKDT	18577
QY	154	-----	153
Db	18578	WGVSSGSSKTKLKIPHLQKCEYVFRVRAENKIGVGPPLDSTPTVAKHKFSPSPGKP	18637
QY	154	-----	153
Db	18638	VVTDITENATVSWTLPKSDGSPITGYMERREVTGKVVVNKTPIADLKFRVTGLIEG	18697
QY	154	-----	153
Db	18698	NTYEFVFAENLAGLSKSPSSDPIKACRPIKPPGPPINPKLKDKSRETADLVWTKPLSD	18757
QY	154	-----	153
Db	18758	GGPILGYVVECOKPGTAQWRINKDELIRQCAFVPGLEGNEYFRIKAANIYEGEP	18817
QY	154	-----	153
Db	18818	RELAESVIAKDILHPPPEVELDVTCDVITVRVGQTIIRLARVKRPEPDITWTKGKVLV	18877
QY	154	-----	153
Db	18878	REKRVLIQDLPRVELQKEAVRADHGKYIISAKNSSGHAQSAIYVNLDRPGPCQNLKV	18937
QY	154	-----	153
Db	18938	TNVTKENCTISWENPLDNGSGSEITNFIVEYRKNQKGWSIVASDVTKRLIKANLANNEY	18997
QY	154	-----	153
Db	18998	YFRVCAENKVGVTIETKTPIILAINIDRPGEPENLHIADKGKTEVYLKRRRDPYDGS	19057
QY	154	-----	153
Db	19058	PNLSYHVERRLKGSDDWVRVHKGSIKETHYMWVDRVCVENQIYEFVRQTKNEGGEWVKTE	19117
QY	154	-----	153
Db	19118	EVVVKEDLQVPVLDLKLGLVTVKAGDTIRLEAGVRGKPPPEVAWTKDKDATDLTRSPRV	19177
QY	154	-----	153
Db	19178	KIDTRADSSKFSLTAKRSDGGKYVVVTAINTAGSFVAYATVNVLDKPGPVRLKIVDVSS	19237
QY	154	-----	153
Db	19238	DRCTVCWDPDDGGCEIQNYILEKCTKRWVWSTYSATVLPFGTIVTRLIEGNEYIFRV	19297
QY	154	-----	153

Db	19298	RAENKIGTGPTTESKPVIAKTKYDKPCRPDPPEVTKVSKEEMTVVWNPPEYDGGKSITGY	19357
QY	154	-----	153
Db	19358	FLEKKEKSHTRWVPVVKSAIPERRMKVQNLLPDHEYQFRVKAENEIGIGEPSLSPRPVA	19417
QY	154	-----	153
Db	19418	KDPIEPGPPPTNPRVVDTTKHSTITLWGKRPVYDGGAPIIGYVVMRPKIADASPDEGWR	19477
QY	154	-----	153
Db	19478	CNAQAOLVRKEFTVTSILDENQEYEFVCAQONQVIGRPAELKEAIKPKLEPEIDLDA	19537
QY	154	-----	153
Db	19538	SMRKLIVRAGCPIRLFAIVRGRPAKVTWRKVGIDNVVRKGQVLDVDTMAFLVIPNSTR	19597
QY	154	-----	153
Db	19598	DDSGKSLTLVNPAGEKAVFVNVRLDTPGPVSDLKVSQVTKTSHVSWAPPENDGGSQV	19657
QY	154	-----	153
Db	19658	THYIVEKREADRKTWSTVTPVEVKKTSFHVNTLVPGNEYIFRVTAVNEYGPGVTDVPKPV	19717
QY	154	-----	153
Db	19718	LASDPLSEDPDRKLEVTMTKNSATLAWLPLRDGGAKIDGYITSYREEEQPADRWTEY	19777
QY	154	-----	153
Db	19778	SVVKDSLVLVTGLKEGKKYKFRVAARNVGVSLPREAGVYEAKEQLLPPKILMPQITI	19837
QY	154	-----	153
Db	19838	KAGKLRIEAHVYGKPHPTCKKKKGEDEVVTSHLAVHKADSSSILIKDVTBKDSGYIS	19897
QY	154	-----	153
Db	19898	LTAENSSGTDQIKIVVMDAPGPPQPPFDISDIDACSLSWHIPLEGGSNITNYIVE	19957
QY	154	-----	153
Db	19958	KCDVSRGDWVTALASVTKTSCRVGKLIPOGEYIFRVAENRFGISEPLTSPKMWQPPFG	20017
QY	154	-----	153
Db	20018	VPSEPKNARVTKVKNDCIFVAMDRPDSGGSPIIGYLIERKERNLLWVKANDTLVRSTE	20077
QY	154	-----	153
Db	20078	YFCAGLVEGLEYSFRIYALNKAGSPSPKPTTEYVTARMPVDPGKPEVIDVTKSTVSLIW	20137
QY	154	-----	153
Db	20138	ARPKHDGGSIIIGYFVBACKLPDGKWKVRCNTAPHQIQEETATGLEKAQYQFRAIART	20197
QY	154	-----	153
Db	20198	AVNISPPSPDPVTILAENVPPIDLSVAMKSLTVKAGTNVCLDQVFGKPMPTVSMK	20257
QY	154	-----	153
Db	20258	KDGTLLKPAEGIKMAMORNCTLELFSVNRKDSGDYTTITAENSSGSKSATIKLVLDKPG	20317
QY	154	-----	153
Db	20318	PPASVINKWYSDRAMLSWEPLEGGSEITNYIVDKRETSRPNWAQVSATVPIITSCSVE	20377
QY	154	-----	153

Db 15998 DYKLRVSAVNAAGEPPGETQPVVAEPQEPFAVELDVSKGGIQIMAGKTLRIPAVVTG 16057
QY 133 ----- 132
Db 16058 RPVPTKVTKEGELDKRWIDNVGTSKELIKDALRKDHGRVITATNSCGSKFAAAR 16117
QY 133 ----- 132
Db 16118 VEVFVPGPVLDPVVTNRKMLNWSDEDDGSEITGFIERKAKMHTWQPIETE 16177
QY 133 ----- 132
Db 16178 RSKCDITLLEGOEYKERVIAKNKFCGPPVEIGPILAVDPLGPPTSPERLYTERTKST 16237
QY 133 ----- 132
Db 16238 ITLDMKEPRSGSGPIOGYIEKRRHDKPDFERVNKRCLPTTSFLVENLDEHQMYEFVK 16297
QY 133 ----- 132
Db 16298 AVNEIGESEPSPLNVVIOQDEVPPTIKRLSVRGDTIKVKAGEPVHIPADVTLGPMPKI 16357
QY 133 ----- 132
Db 16358 EWSKNETVIEKPTDALQITKEVSRSEAKTELSIPKAVREDKGTYYVTASNRLGSVFRNV 16417
QY 133 ----- 132
Db 16418 HVEYDRPSPRNLAVTDIAESCYLTWDAPLDNGSGSEITHVIDKRDASRKAWEVYT 16477
QY 133 ----- 132
Db 16478 NTAVERRYGIWKLIPNGOYEFVRVAVNKYGISDECKSDRWVIOQDPYRLPGLPKPKVLAR 16537
QY 133 ----- 132
Db 16538 TKGSMVSWTTPPLDNGGSPITGWLEKREEGSPYRSVRAPITKVGLKGVEFNVPRLLE 16597
QY 133 ----- PPSR 136
Db 16598 GVKYQFRAMAINAAGIGPPSEPDVAGDPIPPGPPSCPEVKDKTKSSISLGNKPPAK 16657
QY 137 D----- 137
Db 16658 DGGSPIKGYIVEMQEGTDMKRVNEPKLITTCWCVPNLKELRYFRKAVNEAGES 16717
QY 138 ----- 137
Db 16718 EPSDTTGEIPATDIOEPEVFFIDGAQDCLVCKAGSQIRIPAVIKGRPTPKSSWEPDGKA 16777
QY 138 ----- 137
Db 16778 KRAMKDVHDIPEDAQLETAENSSVIIPECKRSHTGKYSITAKNKGOKTANCRVKVND 16837
QY 138 ----- 137
Db 16838 VFGPKDLKVSDITRGSCLSNKMPDDGGDRIKGYIEKRTIDGKAWTKVNPDCGSTTF 16897
QY 138 ----- 137
Db 16898 VVPDLLSEQQYFVRVAENRFGIGPPVETIQRTTARDPIYPPDPKIKGLITKNTVHL 16957
QY 138 ----- 137
Db 16958 SWKPPKNDGSGPVTHYIVECLAWDPTGKKEAWRQCNKRDEVELOFTVEDLVEGGEYFR 17017
QY 138 ----- 137
Db 17018 VRVNAAGVSKPSATVGPCDCORPMPPSIDLKEFMEVEEGTNVIVAKIKGVPPFTLW 17077
QY 138 ----- 137
Db 17078 FRAPPKKPNKEBVLVDYTHVNLVDDTCTLVIPOSRRSDTGLYITITAVNNLGTASKEMR 17137

QY 138 ----- 137
Db 17138 LNVLRGPPVGPFIKESVSADQMTLSWFPKDDGSKITNVYIEKREANRKTWVHSSE 17197
QY 138 ----- 137
Db 17198 PKECTYTIKPLLEGHEVFRIMAQNKYIGCEPLDSEPTARNLFSVGPADKPTVSSVTR 17257
QY 138 ----- 137
Db 17258 NSMTVNWEEPEYDGGSPVTGYWLEMKDITTSKRWRVNRDPIKAMTLGVSKYKVTGLLEGSD 17317
QY 138 ----- 137
Db 17318 YQFRVYAINAAGVGPASLPDPATARDPIAPGPPPKVTDWTKSSADLEWSPPLKDGGG 17377
QY 138 ----- 137
Db 17378 KVTGYIVEYKEEGKEWEKDKVEVGTCLVVTGLKEGAFYKFRVSAVNIAGIGEPVET 17437
QY 138 ----- 137
Db 17438 DVIEMKDRLVSPDLQLDASVRDRIVVHAGGVIRIIAYVSGKPPPTVTWNNERTLPQEA 17497
QY 138 ----- 137
Db 17498 IETTAISSMWIKNCORSHQGVYSLAKNEAGERKKTIIIVDLVDVPGVGTPTPLAHNLIN 17557
QY 138 ----- 137
Db 17558 ESCKLTWFSPEDDGGSPITNYVIEKRESRRAWTPVTYVTRQNAFVQGLIOGKAYFRI 17617
QY 138 ----- ELTKNOVSLT- ----- 147
Db 17618 AAENSIGMGPVETSEALVIREPIVPERPEDLEWKEVTNTVTUTWNPVKYDGSSEIIN 17677
QY 148 ----- : : : : : ----- 147
Db 17678 YVLESRLIGTEKFKHVTNDNLLSRKVTYKLGKEDGYEYRSVAVNIVGQKPSFCFKPIT 17737
QY 148 ----- ----- 147
Db 17738 CKDELAPTLHLDPRDKLITRVGEAFALTCRYSGKPKKVSFWFKDEADVLEDDRTHIKTT 17797
QY 148 ----- CLV- ----- KGF- ----- 153
Db 17798 PATLALEKIKAKRSDSGKYCVVVENSTGSRKGCQVNVVDRPGPPVGPVSFDEVTKDYMV 17857
QY 154 ----- : : : : : ----- 153
Db 17858 ISWKPPLDGGSKITNYIIIEKKEVGKDVWMPVTSASAKTTCVSKLLEKGDYIFRIHAEN 17917
QY 154 ----- ----- 153
Db 17918 LYGISDPLVDSMKAKDRFVPDAPDQPIVTEVTKDSALVTWKNPHDGKPTITNYILEKR 17977
QY 154 ----- ----- 153
Db 17978 ETMSKRWARVTKDPIHTPYTKFRVPDLLECCQYEFVRSAENEIGIDPSPKPVFAKDPI 18037
QY 154 ----- ----- 153
Db 18038 AKPSPVPNEAIDTTCNSVDLTWQPPRHGDSKILGYIVEYOKVGEWRRANHTPESC 18097
QY 154 ----- ----- 153
Db 18098 ETKYKVTGLRDGQTYKFRVLAVNAAGESDPAHVPEVLVKDRLEPELILDANMAREQHI 18157
QY 154 ----- ----- 153
Db 18158 KYGDTLRLSAILIKGVPPFKVTKWKEDRDAPTAKRIDVTPVSGSKLEIRNAAHEDGGIYSLT 18217

QY 133 ----- 132
Db 13838 PVEIIRPPQDILEAPGADVFLAELNKKDKVEVQWLRNNMVMVVOGDHOMMSEGIHRLOI 13897
QY 133 ----- 132
Db 13898 CDIKPRDQGEYRTAKDKARAKLAAAPKIKTADQDLVVDVGKPLTMVVPYDAYPKAE 13957
QY 133 ----- 132
Db 13958 AEWKENEPLSTKIDTTABQTSFRILEAKKGDGKRYKIVLQNHGKAEGFINLKVIDVP 14017
QY 133 ----- 132
Db 14018 GPNRLEVTEFDGEVSLAWEELTDGGSKIIGYVVVERRDIKRTWVLATDRAESCEFTV 14077
QY 133 ----- 132
Db 14078 TGLQKGGVEYLFRVSARNRVGTGEPVETONPVEARSKYDVPGLNVTITDVNRFGVSLT 14137
QY 133 ----- 132
Db 14138 WEPPEYDGAETNYVIELRDKTSIRWDTAMTVRAEDLSATVTVVVEGOEYSFRVAQNR 14197
QY 133 ----- 132
Db 14198 IGVKPSAATPFVKVADPIERPSPVNLTSDDQTQSSVOLKWEPLKDGSPILGIYIER 14257
QY 133 ----- 132
Db 14258 CEBGKDNWIRCNKMLVPELTYKVTGLEKGNKLYRVAENKAGVSDPSEILGPLTADDAF 14317
QY 133 ----- 132
Db 14318 VEPTMDSAPKOGLEVIVPNPITILVPSTGYPRPTATWCFGDKVLETGDRVKMTLSAYA 14377
QY 133 ----- 132
Db 14378 ELVISPERSDKGIYTLKLENRYKTIISGEIDVNVIAEPSAPKELKFGDITKDSVHLTWEP 14437
QY 133 ----- 132
Db 14438 PDDGGSPLTYGVVEKREYSRKTWTKVMDVFTDLEFTVPLVQGEYLFKVCARNKCGPG 14497
QY 133 ----- 132
Db 14498 EPAYVDEPNMSTPATVPDPPENVKWRDRTANSIFLTWDPKNDGGSRIKGYIVERCPRG 14557
QY 133 ----- 132
Db 14558 SDKWACGEPVAETKMEVTGLEEGKWYAVRVKALNRQGASKPSRPTTEEIQAVDTQEAPEI 14617
QY 133 ----- 132
Db 14618 FLDVKILLAGLTVKAGTKIELPATVTKGPEKITWTAKMILKODKRITTIENPKKSTVTI 14677
QY 133 ----- 132
Db 14678 VDSKRSDTGTYYIEAVNVCGRATAVVEVNVLDKPGPPAAFDITDVNESCLLTWNPPRDD 14737
QY 133 ----- 132
Db 14738 GGSKITNYVVERRATDSEVWHKLSSTVKDTNFKATKLIPNKEYIFRVAENMYGVEPVQ 14797
QY 133 ----- 132
Db 14798 ASPITAKYQDPGPPTRLEPSDITKDAVTLTWCEPDDGGSPIGYWVERLDPDQKW 14857
QY 133 ----- 132
Db 14858 RCNKMVPKDTYRVKGLTNKKYFRVLAENLAGCPKPSKSTEPILIKDIPDPWPPGKP 14917
QY 133 ----- 132

Db 14918 TVKDVGKTSVRLNWKTRPEHGDGAKIESYVIEMLKTGTDWVRVAEGVPTTQHLLPGLMEG 14977
QY 133 ----- 132
Db 14978 QEYSFRVRAVNKAGESEPSDPVLCREKLYPPSPRMLVINITKNTADLAKWTVPKED 15037
QY 133 ----- 132
Db 15038 GGSPIITNYIVEKRDVRRKGMQTVDTTVKDKCTVPTLPTEGSLYVFRVAENAIQGSYTE 15097
QY 133 ----- 132
Db 15098 IEDSVLAKDTFTTPGPPYALAVVDVTKRHVDLAKWEPKNDGGRPIQRIYVIEKKERLGTWR 15157
QY 133 ----- 132
Db 15158 VKAGTAGPCDCNFRVTDVIEGTEVQFQVRAENEAGVCHPSEPTIELSIEDPTSPSPPLD 15217
QY 133 ----- 132
Db 15218 LHVTDAGRKHIAIAWKPEKNGGSPIGYHVEMCPVCTEKKMRVNSRPIDKLFKVEEGV 15277
QY 133 ----- 132
Db 15278 VPKKEYVLRVAVNAIGVSEPSISENVVAKDPDCKPTIDLETHDIIIVIEGKLSIPVPF 15337
QY 133 ----- 132
Db 15338 RAVPVPTVSMHKGKVKASDRLTMKNDHISAHLEVPKSVRADAGIYVITLENKLSATA 15397
QY 133 ----- 132
Db 15398 SINVKVIGLPGPKDIAKSDITKSSCKLTWEPPEFGGTPILHYVLERREAGRTYIPVM 15457
QY 133 ----- 132
Db 15458 SGENKLSWTVKDLIPNGEYFRVKAVKVGGEYIELKNPVIADPKQPPDPVDVEVHN 15517
QY 133 ----- 132
Db 15518 PTAEAMTITWKPLYDGGSKIMGYIIETIAKEERWKRCNEHLVPLTYTAKGLEEGKEY 15577
QY 133 ----- 132
Db 15578 QFRVRAENAGISEPSRATPPTKAVDPIDAPKVLRTSLEVKRGDEIADASISGSPYPT 15637
QY 133 ----- 132
Db 15638 ITWIKDENVIVPEIKKRAAPLVRRRKGEVEEPFVPLTQRLSIDNSKKGSQRLVRD 15697
QY 133 ----- 132
Db 15698 SURPDHGLYMIKVENDHGIKAPCTVSVLDTPGPPINFVFEDIRKTSVLCWKPEPLDDGG 15757
QY 133 ----- 132
Db 15758 SEIINTLEKKDKTKPDSEWIVVTSTLRHCKYSVTKLIEGKEYLFRVRAENRFGPGPCV 15817
QY 133 ----- 132
Db 15818 SKPLVAKDPGPDAPDKPIVEDVTSNMLVKWNEPKDNGSPLGYWLEKREVNSTHWSR 15877
QY 133 ----- 132
Db 15878 VNKSLNALKANVDGLLEGLTYVFRVCAENAGPGKSPDPKTAHDPISPPGPIPRV 15937
QY 133 ----- 132
Db 15938 TDTSSTTIELEWEPFAPNGSGGEIVGYVDKQLVGTNENSRCTEKMVKRQYTVKREGA 15997
QY 133 ----- 132

QY 107 ----- 106
Db 9458 NEHEIESNVLQVDERKKQEKIEGDLRALMKKTPILKKGAGEEEDIDIMELLKNVDPKE 9517
QY 107 ----- 106
Db 9518 YEKYARMYGITDFRGLIQAPELLKQSQEEETHRLEIEIERSEDERKEFEELVSFIQORL 9577
QY 107 ----- 106
Db 9578 SQTEPVTLIKDIENQTVLKNDAVFEIDIKINYPEIKLSWKYKTEKLEPSDKFEISIDGD 9637
QY 107 ----- 106
Db 9638 RHTLRVNCOLKDOGNRYLVCGPHIASAKLTVIEPAWHERLQDVTLKEGOTCTMTQCFVS 9697
QY 107 ----- 106
Db 9698 PNVKSEWFRNGRILKPOGRHKTEVEHKVHKLTIAADVRAEDQGYTCKYEDLETSaelRIE 9757
QY 107 ----- 106
Db 9758 AEPIQFTKRIQNIWSEHQSAFCEVSFDDAIVTWYKGTTELTSQKYNFRNDGRCHYM 9817
QY 107 ----- 106
Db 9818 TIHNTPDDEGVSVIARLEPRGEARSTAEIYLTKEIKLELKPDPIDSRVPIPTMPIR 9877
QY 107 ----- 106
Db 9878 AVPEEIPVAPPILPLLTPEEKPPPKRIEVTKKAKKDAKKVAKPKEMTPREEIV 9937
QY 107 ----- 106
Db 9938 KKPPPTTLIPAKAPEIIDVSSKAEEVKIMTITRKKEVQKEAIVYKQAVHKEKRVFI 9997
QY 107 ----- 106
Db 9998 ESPEEPYDELEVFPYEPFQPYEERPEDEYEEIKVEAKKEVHEEEDFEQCQYERE 10057
QY 107 ----- 106
Db 10058 EGYDEGEWEAEAYQEREVIQVQKEVVEESHERRKVPKAPVEKKAPPPKVIKKPVIEKIE 10117
QY 107 ----- 106
Db 10118 KTSRMEEEKVQVTKVPEVSKKIVPQKPSRTPVOEEVIEVKVPVHTKKMVISSEKMFFA 10177
QY 107 ----- 106
Db 10178 SHTEEEVSVTPEVQKEIVTEEKIHVAVSKRVPPPKVPPELPEKPAPEEAVPVPKPKVE 10237
QY 107 ----- 106
Db 10238 PPAPKVPVPPKVPPEKKPVPVPKKEPAAPPKVPVPPKVPPEKIPVPVAKKKEAPPA 10297
QY 107 ----- 106
Db 10298 KVPVQKRVVTEBKITIVTQREESPPPAVPEIPKKVPBEKVPVPRKEEVPVPPKVPAL 10357
QY 107 ----- 106
Db 10358 PKKVPPEKVAVPVPVAKKAPPRAEVSCKTVVVEKKRFVAEEKLSFAVQORVEVTRHEVS 10417
QY 107 ----- 106
Db 10418 ABEWSYSEEGSVISVYREEREBEEAEVTEYEVMBEPEEYVVEEKLHIISKRVAE 10477
QY 107 ----- 113
Db 10478 PAEVTERQEKKIVLPKIPAKIEBPPPAKVPAPKKIVPEKKVPAPVPPKKEKVPVPPKVPPE 10537
QY 114 ----- 113

Db 10538 EPKKVPPEKKVPVKVIMBEEPLAKVTERHMQITQEEKVLVAVTKKEAPPKARVPPEPKR 10597
QY 114 ----- 113
Db 10598 AVPEEKVLKLKPREBEPKAKVTEFRKRVVKEEKVSIAPKREPQPIKEVTIMEEKERAY 10657
QY 114 ----- 113
Db 10658 TLEEAHSVQREBEYEEYEDYKFEFEYEPTEYDQYEEYERERYEHEEYITEPE 10717
QY 114 ----- 113
Db 10718 KPIPVKVPPEPVTPKPAKPAKVLKKAPEEKVPVPIPKKLKPPPKVPPEKKVPPEK 10777
QY 114 ----- 113
Db 10778 IRISITKREKEQVTEPAAKVPMKPRVVAEEKVPVPRKEVAPVPRVPEVKELEPEEVAF 10837
QY 114 ----- 113
Db 10838 EEEVTHVEEYLVVEEYIHEEEFITEEEVVPVTPVKVPEVPRKVPPEKKVPVPPKK 10897
QY 114 ----- 113
Db 10898 KEAPPKVPPEKKVPPEKVPVLIPKKEKPPPAKVPPEKVPPEKVPVPPKKEAPPA 10957
QY 114 ----- 113
Db 10958 KVPPEKKVPPEKKVPVPAKVPKVEAPPKVPPEKVLIPPEKKTPVPKKEAPPKVPK 11017
QY 114 ----- 113
Db 11018 KREPVVPVALPOEEVLFEEIIVPEEEVLPPEEEVLPPEEEVLPPEEEIPE 11077
QY 114 ----- 113
Db 11078 EEEVPEEEVPEEEFVPEEEVLPVKVPVPAVPVPEIKKKVTKKKVLPKKEAPPA 11137
QY 114 ----- 113
Db 11138 KVPPEKKVVEEKRIILPKKEEVLPEVTERPEEPISEEEIPEEPPSIEEVEEAVPRVP 11197
QY 114 ----- 113
Db 11198 EVIKKAVPEAPTVPKKEAPPKAKVSKKIPEEKVPVPOKKEAPPKVPPEKVPPEKKV 11257
QY 114 ----- 113
Db 11258 LVPKKEAVPPAKGRTVLEEKVSVAFRQEVVVKERLELEVVEAEVEEIPPEEEFHEVEYF 11317
QY 114 ----- 113
Db 11318 EEEGHEVEEYFIKLEQHRVEEHEHREKVRHREVEFAEEVEVEFKPAKPKGPEISEKII 11377
QY 114 ----- 113
Db 11378 PPKKPTKVVPRKEPPKVPPEKVIWEEKVRPEEPRVPTKVPDVLPPKEVPEKKV 11437
QY 114 ----- 113
Db 11438 PVPPAKKPEAPPKVPPEKVPPEKKVPVPPPKPEVPPTKVPPEKAAVPEKKVPEAI 11497
QY 114 ----- 113
Db 11498 PPKPESEPPPEVPPEAPKEVVPPEKKVPAAPKPKPEVTPVKPEAPKEVVPPEKKVPPPKP 11557
QY 114 ----- 113
Db 11558 EVPTTKVPEVPKVPKVPPEKKVPEAIPPKPSPPPEVEEPEEVALEEPBAEVEEPEPAAP 11617
QY 114 ----- 113

Db 7238 IRPGNYTITCVGNTPHLRILKVGKDSGQYTCQATNDVGKDMCSAQLSVKRPFRVKKL 7297
QY 90 -----LHQDW----- 94
Db 7298 EASKVAKOGESIQLECKISGPEIKVSWFRNDSELHESWKYNMFSINVSALLTINEASAE 7357
QY 95 -----||: |----- 94
Db 7358 DSGDYICEAHNGVGASCSTALTIVKAPVFTQKPSVGALKGSDVILQCEISGTPPFVW 7417
QY 95 ----- 94
Db 7418 WVKDRKQVRNKKFKITSKHFDTSLHILNEASDVGEYHCKATNEVGSDTSCSVKFKPE 7477
QY 95 -----||: |----- 94
Db 7478 PRFVKKLSDTSLIGDAVELRAIVEGFQPSVWLKDRGEVIRESENTRISFIDNIATLQ 7537
QY 95 ----- 94
Db 7538 LGSPEASNGKYICQIKNDAGMRECSAVLTVLEPARIIEKPEPMVTVTGTFALCEVVTG 7597
QY 95 ----- 94
Db 7598 TPELSAKWFKDGRSADSCKHHITFINKVASLKIPCAEMSKGLYSFEVKNVSGKSNCTV 7657
QY 95 ----- 94
Db 7658 SVHVSDRIVPPSFIRKLKDVNALCASVVLCEVRVSGSAPISVGWFQDGNIEVSGPKQSS 7717
QY 95 ----- 94
Db 7718 FSENVCTLNLSLEPDTGIYTCVAANVAGSDECSAVLTVOEPPSFQTPDSVEVLPGMS 7777
QY 95 ----- 94
Db 7778 LFTFSVIRGTPPFVKWFKGSRVLPGESCNISLEDFVTELELFEVQPLESGDYSLVTN 7837
QY 95 ----- 94
Db 7838 DAGSACTHLFVKEPATVFKRLADFVETGSPVILEATYTGTPISVSWIKDEYLIQS 7897
QY 95 ----- 94
Db 7898 ERCSTMTKSTILEILESTIEDYAYQSCLINEAGODICEALVSLVLEPPYFIEPLEHVE 7957
QY 95 ----- 94
Db 7958 AVIGEPATLOCKVDGTPETIRISWYKEHTKLSAPAYKMQFKNNVASLVINKVDHSDVGEY 8017
QY 95 -----LNGKE----- 99
Db 8018 SCKADNSVCAVASSAVLVIKERKLPPFFARKLKVHETLGFVPAECRINGSEPLQVSWY 8077
QY 100 -----||: |----- 99
Db 8078 KGVLLKDDANLQTSFVHNATLQILOQDSHIGQYNCSASNPLGTASSAKLILSEHEV 8137
QY 100 ----- 99
Db 8138 PPFDDLKPVSDVLALGESGTFKCHVTGTAPKITWAKDNREIRPGNGYKMTLVENTATLT 8197
QY 100 ----- 99
Db 8198 VLKVGKGDAQYTCYASNLAGKSCSAHLGVQEPFRFICKLEPSRIVKQDETRYECKIG 8257
QY 100 ----- 99
Db 8258 GSPEIKVLWKDETEIQESSKFRMSFVDSVAVLEMHNLSVEDSGDYTCEAHNAGSASS 8317
QY 100 ----- 99
Db 8318 TSLKVKEPPIFRKKPHPIETLKGADVHLECELGQTPPFPHSVWYKDKRELRSCKYKIMSE 8377

QY 100 ----- 99
Db 8378 NFLTSIHILNVDAADIGEYQCKATNDVGSCTCVGSIALKAPPRFVKKLSLSDISTVVGKEVQ 8437
QY 100 ----- 99
Db 8438 LQTTIEGAEPISVWVFKDKGEIVRESNDIWIISYSENIATLQFSRVEPANAGKYTCQIKND 8497
QY 100 ----- 99
Db 8498 AQMQECFATLSVLEPATIVERPESIKVTTGDTCTLECTVAGTPELSTKFWKDKGKELTSDN 8557
QY 100 ----- 99
Db 8558 KYKISFFKNVSGKLIINVAPSDSGVYSFEVQNPVGKDSCTASLQVSDRTVPSPFTRKLKE 8617
QY 100 ----- 99
Db 8618 TNLGSSVWMECKYVGGPPISSVWFHEGNEISSGRKYQTTLTNTCALTVMNLEESDGG 8677
QY 100 ----- 99
Db 8678 DYTCTIATNWAGSDECSAPLTVREPPSFVQKPDPMVDLGTNTVFTTSIVKGTPPFSVSWFK 8737
QY 100 ----- 99
Db 8738 GSSELVPGDRCNVLSLEDSVAELELFDVDTQSQGEYTCIVSNEAGKASCTTHLY IKAPAKF 8797
QY 100 ----- 99
Db 8798 VKRLNDYSIEKGKPLILEGTFGTGTPPISVWTKKNGINVTSPQRNITTTESAILLEPSS 8857
QY 100 ----- 99
Db 8858 TVEDAGQNCYENASGKSCSAQILILEPPYFVKQLEPVPKVSVDGSASLQCLAGTPEI 8917
QY 100 ----- 99
Db 8918 GVSWKGTCLRPRTTYTKMHFRNNVATLVFNQVDINDSGEYICKAENSVEVSASTFLTIV 8977
QY 100 ----- 99
Db 8978 QEQKLPPSFRQLRDVQETVGLPVVFDCAISGEPIPSVSWYKDGKPLKDSPNVQTSFLDN 9037
QY 100 ----- 99
Db 9038 TATLNIKTDRSLAGQYSCATNPIGSASSARLILTEGKNPPFFDIRLAPVDVAVGESA 9097
QY 100 ----- 99
Db 9098 DFECHVTGTQPIKVSWAKDSREIRSGGKYQISYLENSAHLTVLKVDKDGSGQYTCYAVNE 9157
QY 100 ----- 99
Db 9158 VKDSCSTAQLNIKERLIPPSTFKRLSETVEETEGNSFKLEGRVAGSQPITTVAMYNNEI 9217
QY 100 -----YKCKVSN----- 106
Db 9218 QPSTNCEITFRNNTLVLOVRKAGMNDAGLYTCKVSNDAAGSALCTSSIVIKEPKKPPVFDQ 9277
QY 107 -----||: |----- 106
Db 9278 HLTPTVSEGEYVQLSCHVQSGSEPIRIQWLKAGREIKPSDRCSFSFASCTAVLELRDAVK 9337
QY 107 ----- 106
Db 9338 ADSDGYVCKASNAGSDTTKSKVTIKDKPAVAPATKKAADVGRLEFFVSEPIRVVEKTT 9397
QY 107 ----- 106
Db 9398 ATFTIAKVGDDIPNVKWTGKWRQLNQGRVFIHOKGDEAKLEIRDTTKTDSGLRYCVAF 9457

QY	81	-----80
Db	5078	KEIADRYRIAFVEGTASLEIIRVDMNDAGNFTCRATNSVSGSKDSSGALLIVQPPSFVT 5137
QY	81	-----80
Db	5138	KPGSKDVLPGSAVCLAKSTFOGSTPLTIRWPKGNKELVSGGSCYITKEALESLELYIVKT 5197
QY	81	-----80
Db	5198	SDSGTYTCKVSNVAGGVECSANLFEVKEPATFVEKLEPSQLLKKGDATQLACKVTGTPPIK 5257
QY	81	-----80
Db	5258	ITWFANDREIKESKHRMSFVESTAVLRLTDVGIEDSGEYMCFAQNEAGSDHCSSIVIVK 5317
QY	81	-----80
Db	5318	ESPYFTKEPKPIEVLKREYDYMLLAEVAGTTPPEITWFKDNTILRSGRKYKTFIQDHLVSL 5377
QY	81	-----80
Db	5378	QILKFAADAGEYQCRVTNEVGSSICARSVTLREPPSFYKKBESTSSLRGGTAFAOATLK 5437
QY	81	-----80
Db	5438	GSLPIITVTLKDSDEITEDDNIRMTFENNVAASYLSGIEVKHDKYVCQAKNDAGTORCS 5497
QY	81	-----80
Db	5498	ALLSVKAPATITEAVSIDVTQGPATLQVKFSGTKEITAKWFKDQELTLGSKYKISVT 5557
QY	81	-----80
Db	5558	DTVSIILKIISTEKDSGEYTFEVDNGRSCSKARINVLDLIIPSTFKLKMDSIKGS 5617
QY	81	-----80
Db	5618	FIDLECTIVAGSHPISIOWFKDDQEIASEKYKFSFHDNTAFLEISQLEGTDSTYTCSAT 5677
QY	81	-----80
Db	5678	NKAGHNQCSGHLVKEPPYFVEKPSQDVNPTRVOLKALVGTAPTIKWFKDNKELHS 5737
QY	81	-----80
Db	5738	GAARVWKDDTSTSLFAAKATDSGTYYICQLSNDVGTATSKATLFVKEPPQFIKKPSV 5797
QY	81	-----80
Db	5798	LVLNQGOSTTECQITGTPKIRVSWYLDGNEITAIQKHGISFIDGLATFQISGARVENSG 5857
QY	81	-----80
Db	5858	TYVCEARNDAGTASCIELVKKEPPTFIRELKPVEVVKYSDVECEVTGTPPEVTWLK 5917
QY	81	-----80
Db	5918	NNREIRSSKYYLTDRVSVFNHLITKCDPDTGEYQCIIVSNEGGSCSTRVALKEPPSF 5977
QY	81	-----80
Db	5978	IKKIENITVLKSSATFQSTVAGSPISITWLKDDQILDEDDNNVYISFVDSVATLQIRSV 6037
QY	81	-----80
Db	6038	DNGHSGRYTQAKNESGVERCYAFLVLQBPQAIVEKAKSVDVTEKDPMTLECVVAGTPEL 6097
QY	81	-----80
Db	6098	KVKWLKDGKOIVPSRYFSMFENNVAFRISQVMKODSQYTFKVENDFGSSCDAYLRV 6157
QY	81	-----YRVV----84

Db	6158	LDQNIIPPSTKLTLMKMDKVLGSSIIHMECKVSGSLPISAQWFKDCKEISTSAKYRLVCHER 6217
QY	85	-----84
Db	6218	SVSLEVNLEEDTANYTCKVSNVAGDDACSGILT'VKEPPSFLVKPGRQOAI'PDSTVEFK 6277
QY	85	-----84
Db	6278	AILKGT'PPKIKWFKDDVELVSGPKCFIGLEGSTSFNL'YSVDASKTGQY'TCHVTNDVGS 6337
QY	85	-----84
Db	6338	DSCTTMLLVTEPPKFVKKLEASKIVKAGDSSRLECKIAGSPEIRVVMFRNEHEL'PASDKY 6397
QY	85	-----84
Db	6398	RMTFIDSVAVIOMNNLSTEDSGDFICEAQN'PAGTSCST'KVIVKEPPVFSSFPPIVETLK 6457
QY	85	-----84
Db	6458	NAEVSLECELSGTPPFVWVYKDKRQLRSKKYKIAKSNFHTSIHILNVDTSIGEYHCK 6517
QY	85	-----84
Db	6518	AQNEVGSDTCVCTVKLKEPPRFVSKLNSLTVVAGEPAELQASIEGAQPIFVQWLKEKEEV 6577
QY	85	-----84
Db	6578	IRESENIRITFVENVATLOFAKAEPANAGKYICQIKNDGMEENMATLMVLEPAVIVEKA 6637
QY	85	-----84
Db	6638	GPMTVTVGETCTLECKVAGTPELSVENYKDGKLLTSQXHKFSFYKNKISSRLILSV'ERQD 6697
QY	85	-----84
Db	6698	AGTYTFQVQNNVCKSSCTAVVDVSDRAVPPSFTRRLKNTGGVLGASCILLECKVAGSSPIS 6757
QY	85	-----SVLTV--89
Db	6758	VAMFHEKTKIVS'GAKYQTTTFSDNVCT'LQLNSLSDSDMGNYT'CAANVAGSDECR'AVLT'VQ 6817
QY	90	-----89
Db	6818	EPPSFVKEPEPLEVLP'GKNV'TFTSVIRGTPPKVNMFRGARELVKGDR'CNIFYEDTVAEL 6877
QY	90	-----89
Db	6878	ELFNIDISQSGEYTCVVSNAGQASCTTRL'FVKEPA'AF'LKRLSDHSVE'PCKSII'LESTYT 6937
QY	90	-----89
Db	6938	GTLPI'SVTWKDG'FNITSEKCNIVTTEKTCILEILNLT'KRDAGQYSC'EIENAGRDVCG 6997
QY	90	-----89
Db	6998	ALVSTLEPPYFVTELEPLEA'AVGDSV'LCQVAGTPEITVSWYKGT'KL'RPTEYTYFT 7057
QY	90	-----89
Db	7058	NNVATLVFNKVNINDSGEYTC'KAENSIGT'ASSKTVPRI'QERQLPPSFARQLKDIEQTVGL 7117
QY	90	-----89
Db	7118	PVTLTCL'RLNGSAPIQVCWYR'GVDG'VLLRDDENLOT'SFVDNVATILKILQTDLSHSGQYSCSAS 7177
QY	90	-----89
Db	7178	NPLGTASSARLTAREPKKSPFFDIKPVSIDVAGESADFECHVTGAQPMRITWSKDNKE 7237
QY	90	-----89

Db 2858 PSDAGEYTAVGQLECKAKLFVETLHITKMKNIEVPETKTASFCEVSHFNVPSMWLKN 2917
QY 74 ----- 73
Db 2918 GVEIEMSEKFIKVOGKLHQLIIMNTSTEDSAEYTFVCGNDQVSATLVTPTIMTSMKLD 2977
QY 74 ----- 73
Db 2978 INAEKDTITFEVTVNYEGISYKWLKNGVEIKSTDKCQMRKTLTHSLNIRNVHFGDAAD 3037
QY 74 ----- 73
Db 3038 YTFVAGKATSTATLYVEARHIEFRKHIDIKVLEKKRAMFECEVSEPDITVOMKDDQEL 3097
QY 74 ----- 73
Db 3098 QITDRIKIQEKYVHRLIIPSTRMSDAGKYTVVAGNVSTAKLFVEGRDVRIRSIKKEVQ 3157
QY 74 ----- 73
Db 3158 VIEKQAVVEFVNEDDVAHWYKDGIEINFQVQERHKYVVERIRHMFISETROSADAGE 3217
QY 74 ----- BEQ 76
Db 3218 YTFVAGNRSSVTLVNAPEPPQVLQELQPVTVQSGKPARFCNAVISGRPOPKISWYKBEQ 3277
QY 77 YNST----- 80
||
Db 3278 LLSTGFKCKFLHDGQEYTLLEAFPEDAAVYTCENKNDYGWATTSASLSVEVPEVSPD 3337
QY 81 ----- 80
Db 3338 QEMPVYPAIITPLOTVTSEGQPARFQCVSGTDLKVSYSKDKKIKPSRFRMTQPED 3397
QY 81 ----- 80
Db 3398 TYOLEIAEAYPEDEGTYTFVNASNAVQVSSANLSLEAPESILHERIEOEIEMEMKEPFSS 3457
QY 81 ----- 80
Db 3458 SFLSAEEGLHSAELQSLKINETLELLESPPVSTKFDSEKGTGPIFIKEVSNADISMG 3517
QY 81 ----- 80
Db 3518 DVATLSVTVIGIPKPIQWFFNGVLLTPSADYKFVFDGDHSLIILFTKLEDEGEYTCMA 3577
QY 81 ----- 80
Db 3578 SNDYKGTICSAYLKINSKGEHGDTESETSAVAKSLEKLGPCPPHFLKELKPIRCAQGLP 3637
QY 81 ----- 80
Db 3638 AIFEYTVVGEPAPTVTFEKENQLCTSVYTIHNPNGSGTFIVNDQREDSGLYICKAE 3697
QY 81 ----- 80
Db 3698 NMLGESTCAEALLVLEDDMTDTPCKAKSTPEAPDFPQPLKGPAVEALDSEGEIATF 3757
QY 81 ----- 80
Db 3758 VKDTILKAALITEENQOLSYEHIAKANELLSQLPLGAQELQLEODKLTPESTREFLCI 3817
QY 81 ----- 80
Db 3818 NGSIHFQPLKPSNQLQIVOSOKTFSKEGILMPEEPETOAVLSDTEKIFPSAMSIEQI 3877
QY 81 ----- 80
Db 3878 NSLTVEPLKTLAEPEGNYPOSSIEPPMHSYLTSVAEVLSPKKTVDNREQRVTLOK 3937
QY 81 ----- 80
Db 3938 QEAQSALILSQSLAECHVESLQSPDVMISQVNYEPLVPSEHSECTEGGKILIESANPLENA 3997

QY 81 ----- 80
Db 3998 QODSAVRIEEGKSLRFLALEEKQVLLKEHSDNVVMPDQILLESKREPVAIKKVQEVQ 4057
QY 81 ----- 80
Db 4058 RDLLESKESLLSGIPPEQRNLKIQICRALQAAVASQPGFSEWLRNRIEKVEAVNITQ 4117
QY 81 ----- 80
Db 4118 EPRHMCWYLVTSKASVTEVTIIIEDVDPQMANLKMELRDALCAIIEEIDILFAEGPR 4177
QY 81 ----- 80
Db 4178 IQOGAKTSLQEEWDSFSGSQKVEPITEPEVESKYLISPEEVSFYVQSVRVKYLDPVTK 4237
QY 81 ----- 80
Db 4238 GVASAVSDEKODESLKPFSEKESSESSEGTVEEVATVKIQEAEAGGFIKEDGPMIHTPLVD 4297
QY 81 ----- 80
Db 4298 TVSEEGDIVHLTTSITNAKEVNWYFENKLVPSDEKFKLQDQNTYTLVIDKVNTEHDQGE 4357
QY 81 ----- 80
Db 4358 YVCEALNDSGTKTSAKLTVVKRAAPVIKRIKTEPLEVALGHLAKFTCEIQSAPNVRFQWF 4417
QY 81 ----- 80
Db 4418 KAGREIYESDKGSIIRSKYISSLEILRTQVDCGEYTCASNEYGSVSTATLTVTEAYP 4477
QY 81 ----- 80
Db 4478 PTFILSRPKSLTTFVGAAKFICTVTGTPVIETIWOKDGAALSPSPNWKISDAENKHILEL 4537
QY 81 ----- 80
Db 4538 SNLTIDRGVYCSKASNKGADICQAEI I IDKPHFIKELEPVQSAINKKHLEQOVDED 4597
QY 81 ----- 80
Db 4598 RKVTVTWSKGOKLPPGKDYKICFEDKIATLEIPLAKLSDGYVCTASNEAGSSCSAT 4657
QY 81 ----- 80
Db 4658 YTVREPPSFVKVDPYSYLMPLPGESARLHCKLKGSPVIOVTFKNNKELSESNTVRMYFVN 4717
QY 81 ----- 80
Db 4718 SEAILDITDKVEDSGSYSCAEAVNDVGDSCSTEIVIKEPPSFIKTEPADIVRGTNALL 4777
QY 81 ----- 80
Db 4778 QCEVSGTGPFEISWFKDKKQIRSKKYRLFQSKSLVCLEIFSFSADVGEYECVVANEVG 4837
QY 81 ----- 80
Db 4838 KGCMAATHLLKEPPTFVKVDDLIALGGOTVTLQAAVRGSEPISTVWMMKGQEVIREDKGI 4897
QY 81 ----- 80
Db 4898 KMSFNGVAVLIIPDVQISFGGKYTCLEAENAGSOTVSGELIVKEPAKIERAELIQVTA 4957
QY 81 ----- 80
Db 4958 GDPATLEVTVAGTPELKPWKYKGRPLVASKKYRISFKNNVQLKFSYSAELHDSGQYTFE 5017
QY 81 ----- 80
Db 5018 ISNEVSSSCETTFTVLDRDIAFFTFTKPLRNVDSVVVNGTCRLDCKTAGSLPMRVSFWKDG 5077

QY 10 ----- 9
Db 698 KAEAVATVAADQARVREPREGHLEESYAQQTILEYGKERSAAKVAEPQRPASEP 757
QY 10 ----- 9
Db 758 HVPKAVPRVIOAPSETHIKTTDOKGMHISQIKKTTDLTTERLVHVDKRRPTASPHET 817
QY 10 ----- 9
Db 818 VSKISVPKTEHGYEASIAIATLOKELSATSSAOKITKSVKAPTVPKSETRVRAEPTP 877
QY 10 -----CPAP-- 13
Db 878 LPQFPADPTDYKSEAGVEVKVCVSGITGTVREERFEVLHGREAKYTETARVPAPVE 937
QY 14 -----ELGQPSVFL--FP-- 25
Db 938 IPVTPTLVGLKNVTVIEGESVTLCHISGVPSTVTWYREDYQIESIDFOITFQSGI 997
QY 26 ----- 25
Db 998 ARLMIREAFADSGRTCSAVNEAGTVSTCYLAVOVSEEFKETTAVTEKFTTEKRFV 1057
QY 26 -----PKP----- 28
Db 1058 ESRDVVMTDTSLTEQAGPEPAAPYFITKPVVQKLVEGCVVFGCVGNPKPHYWK 1117
QY 29 ----- 28
Db 1118 SGVPLTGYRYKVSYNKQTECKLVISMTFADDAGETIIVVRNKHGETSASASLLEADY 1177
QY 29 -----KDTLM-- 33
Db 1178 ELLMKSQEMLYQTVTAFVQEPKVGCTAPGFVYSEYEKEQEALIRKKMAKOTVVVR 1237
QY 34 ----- 33
Db 1238 TVVEDQEFHISSEERLIKEIVRIKTTLEELLEDEGEKMAVDISESEAVESGFDLRI 1297
QY 34 ----- 33
Db 1298 KNYRILEGMGVTHFCKMSGYLPKIAWKDGRKRIKHGERYQMDFLQDGRASLRIPVYLPE 1357
QY 34 -----ISRTP-- 38
Db 1358 DEGIYAFASNIKGNALCSGKLYVEPAAPLGAPTYIPTLEPVSIRSLSPRSVSRPIRM 1417
QY 39 ----- 38
Db 1418 SPARMSPARMSPARMSPGRRLEETDESQERLYKPVFVLKPVSPKCLEGQTARFD 1477
QY 39 ----- 38
Db 1478 LKVVGRMPETFWHDGOQIVNDYTHKVVIKEDGTOSLIIVPATPSDGEWTVVAQNAG 1537
QY 39 ----- 38
Db 1538 RSSISVILIVEAHEHQVKPMFVEKLNKVNKEGSQLEMKVRATGNPNPDIVMLKNSDIIV 1597
QY 39 ----- 38
Db 1598 PHYPKIRIEGTGEAALKIDSTVSQDSAWYTATAINKAGRDTTRCKVNVEFEFAPEPE 1657
QY 39 ----- 38
Db 1658 RKLIIIPRGTYRAKETAAPLEPLHRYGOEWEGDLYDKKQKPPFKKLTSLRKR 1717
QY 39 -----EUTC 42
Db 1718 GPAHFECRLTPIGDPMTVVVWVWLDHDKPLEAANRLRMINIEFGYCSLDYGVAYSRDSGIIITC 1777
QY 43 ----- 42

Db 1778 RATNKYGTDHTSATLIVKDEKSLVESQLPEGRKGLQRIEELERMAHEGALTCTVTTDQKE 1837
QY 43 ----- 42
Db 1838 KOKPDIPLYPEPVRVLEGETARFCRVTGYPOPKVWYLNQOLIRSKRFRVYDGIHYL 1897
QY 43 ----- 42
Db 1898 DIVDCKSYDTGEVKVTAENPEGVIEHKVLEIQOQREDFRSLVRRAPRPEPFHVHEPGKL 1957
QY 43 ----- 42
Db 1958 QPEVQKVRPVDITETKEVVKLKRAERITHEKVPEESEELRSFKRRTBEGYVEAITAVE 2017
QY 43 ----- 42
Db 2018 LKSRKXDESYEELLRRKTDELHMTKELTEEEKKALAEKGKITPTFKPKIELSPSMEA 2077
QY 43 -----V 43
Db 2078 PKIFERIQSQTVGQGSDAHFRVRVCKPDPECEWYKNGVKIERSDRIYWNPEDNVCELY 2137
QY 44 VDVSHED----- 51
Db 2138 IRDVTAEADSASIMVKAINIAGETSSHAFLLVQAKQLITFTQELQDVVAKEKDTMATPECE 2197
QY 52 ---PEVKFNWYDQGEVH-----NAKT- 70
Db 2198 TSEPFYKWKYKDGMEVHEGDKYRMHSDRKHVHLSILTIDTSDAEDYSCVLVEDENVKTT 2257
QY 71 ----- 70
Db 2258 AKLIVEGAVVEFVKELQDIEVPESYSGELEICVSPENIEGKWYHNDVELKSNCKYITTSR 2317
QY 71 -----KPR----- 73
Db 2318 RGRQNLTVKDVTKEDQGEYSFVIDGKKTCKLKNKPRPAILQGLSDQKVCEDIVQLEV 2377
QY 74 ----- 73
Db 2378 KVSLESVEGVMKQGOEVOQPSDRHVHIVDKQSHMLIEDMTKEDAGNYSFTIPALGLSTS 2437
QY 74 ----- 73
Db 2438 GRVSVYSVDVITPLKDVNVIEGTKAVLECKVSVPDVTSVKWYLNDEQIKPDDRQVAIVKG 2497
QY 74 ----- 73
Db 2498 TKQRLVNRTHASDEGPKLIVGRVETNCNLSVEKIKIIRGLDLCTCTETQNVVFEVELS 2557
QY 74 ----- 73
Db 2558 HSGIDVLWNFKDKEIKPSSKYKIEAHGKIYKLTVLNMMKDDGKYTFYAGENTISGLTV 2617
QY 74 ----- 73
Db 2618 AGGAIKPLTDQTVAESQEAVFCEVANPDSKGEWLBDGKHLPDNNIRSESDGHRRLI 2677
QY 74 ----- 73
Db 2678 IAATKLDDIGEYTKVATSKTSAKLKEAVKIKKTLKNLTVTETQDAVFTVELTHPNVKG 2737
QY 74 ----- 73
Db 2738 VQMIKNGVVLNESKEYAISVKGTIYSLRIKNCIAIVDESIVGFRGLGASARLHVETVKI 2797
QY 74 ----- 73
Db 2798 IKKPKDVTALENATVAFESVSHDTPVVKWFHKNVEIKPSDKHRLVSRKVKHMLQNLIS 2857
QY 74 ----- 73

RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC010327; AAH10327.1; -
 DR MGI; MGI:21444967; AU044919.
 DR InterPro: IPR000345; Cyt_c_heme_bind.
 DR InterPro: IPR000306; Ig_MHC.
 DR Pfam; PF00047; Ig; 3.
 DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
 KW Hypothetical protein.
 SQ SEQUENCE 473 AA; 51946 MW; CF625F008932AF12 CRC64;
 Query Match 59.0%; Score 787; DB 11; Length 473;
 Best Local Similarity 50.0%; Pred. No. 3.3e-23;
 Matches 143; Conservative 34; Mismatches 47; Indels 62; Gaps 2;
 QY 5 HT-----CP 8
 DB 188 HTPALLQSLGTYMSSVTPSPSTVTCVAHPASSTVVDKKLEPGSTINPCP 247
 QY 9 P-----CPAPELLGGPSVFLPPPKDLMISRTPEVTCVVVDVSHEDPEVKFNVDG 62
 DB 248 PCKECHCKAPNLEGSPSFIFFPNKLDVLMISLTPKTCVVVDVSDPDVDVQISFVNN 307
 QY 63 VEYHNAKTPREQYNSTYRVSVLVVGLVHQLDMLNGKEYKCKVSKNALKAPAPIERTISKAG 122
 DB 308 VEVHTAQTOHREDYNSTIRVVSALPQHODMWSGKEFKCKVKNKDLPSPIERTISKIG 367
 QY 123 QPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLDSD 182
 DB 368 LVRAPOVYTLPPAPQSLRQKDVSLTCLVGFNPGDISVETWSNGHTEENYKDTAPVLDSD 427
 QY 183 GSFFLYSKLTVDSRWQGVNFVCSVMHEALHNHYTOKLSLSPGK 228
 DB 428 GSYFYISKLDIKTSKEKTDSPSCVNRHEGLKNYILKLIISPGK 473
 RESULT 13
 Q8WZ42 PRELIMINARY; PRT; 34350 AA.
 AC Q8WZ42;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Titin.
 GN TTN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 ON NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20309627; PubMed=10850961;
 RA Freiburg A., Trombitas K., Hell W., Cazorla O., Fougereousse F.,
 RA Centner T., Kolmerer B., Witt C., Beckmann J.S., Gregorio C.C.,
 RA Granzier H., Labeit S.;
 RT "Series of exon-skipping events in the elastic spring region of titin
 RT as the structural basis for myofibrillar elastic diversity.";
 RL Circ. Res. 86:1114-1121(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21573839; PubMed=11717165;
 RA Bang M.L., Centner T., Fornoff F., Geach A.J., Gotthardt M.,
 RA McNabb M., Witt C.C., Labeit D., Gregorio C.C., Granzier H.,
 RA Labeit S.;
 RT "The complete gene sequence of titin, expression of an unusual -700
 RT 'Aa titin isoform and its interaction with obscurin identify a novel
 RT 2-line to 1-band linking system.";
 RL Circ. Res. 89:1065-1072(2001).
 DR EMBL; AJ277892; CAD12456.1; -
 DR InterPro: IPR000282; Cytok_receptor_2.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR000577; FGGY_kin.

DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR001092; HLH_basic.
 DR InterPro: IPR003599; Ig.
 DR InterPro: IPR003598; Ig_c2.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003596; Ig_v.
 DR InterPro: IPR002016; Peroxidase.
 DR InterPro: IPR004168; PPAK_motif.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam; PF00041; fn3; 132.
 DR Pfam; PF00047; Ig; 146.
 DR Pfam; PF00069; pkinase; 1.
 DR Pfam; PF02818; PPAK; 53.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00060; FN3; 133.
 DR SMART; SM00409; IG; 167.
 DR SMART; SM00408; IGC2; 148.
 DR SMART; SM00406; IGV; 23.
 DR SMART; SM00220; S_TKC; 1.
 DR SMART; SM00219; TYKc; 1.
 DR PROSITE; PS00933; FGGY_KINASES_1; UNKNOWN_1.
 DR PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
 DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
 DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
 SQ SEQUENCE 34350 AA; 3816262 MW; 5B1120058A7CE58A CRC64;
 Query Match 36.4%; Score 486; DB 4; Length 34350;
 Best Local Similarity 0.5%; Pred. No. 9.7e-05;
 Matches 160; Conservative 43; Mismatches 40; Indels 33574; Gaps 45;
 QY 3 KTHTCPP-----9
 DB 98 KAETAPPNFVRLQSMVTROGQVRLQVRVTGTPVVKFYRDGAETQSSLDFOISEGD 157
 QY 10 -----9
 DB 158 LYSLLIAEAYPEDSGIYVSNATNSVGRATSTAELLVQGEVEPAVKTKTIVSTAQISESR 217
 QY 10 -----9
 DB 218 QTRIEKKIEAHFDARSATVEMVIDGAAGQOLPHKTPHRIIPKPKSRPTPPSIAKAQL 277
 QY 10 -----9
 DB 278 AROQSPSPTRHSPSPVHRVRAPTSPVRSVSPAARISTSPRSVSPLLMRKTOASTVAT 337
 QY 10 -----9
 DB 338 GPEVPPPKQEGYVASSSEAEEMRETTLTSTQIRTEERWEGYGVQVQVTSIGAAGAAAS 397
 QY 10 -----9
 DB 398 VSASASYAAEAVATGAKEVKQDADKSAAVATVVAADVMAVRVPSAVEQTAQRTTTTA 457
 QY 10 -----9
 DB 458 VHIQPAQEQVRKEAEKTAIVTKVVVAADKAKEQELKSRTEVITTKQEQMHVTHQIRKET 517
 QY 10 -----9
 DB 518 EKTFFPKVVISAAKAEQETRISEITKKQKVQTQEAIRQETETAASWVVVATAKSTKL 577
 QY 10 -----9
 DB 578 ETVPGAQEETTQDDQMHLSYEKIMKTRKTVVVPKIVATPKVKEQDLVSRGREGITTKR 637
 QY 10 -----9
 DB 638 EQVQITQEKMRKEAKETALSTIAVATAKAEQETILTRETMTATRQEQIQVTHGKVDVGK 697


```
Query Match      68.5%; Score 914; DB 6; Length 337;
Best Local Similarity 56.8%; Pred. No. 6.6e-29;
Matches 163; Conservative 33; Mismatches 29; Indels 63; Gaps 2;

QY 5 HT-----6
DB 51 HTFSPVLQSSGFYSLSSMVTVPASTWTSETYICNVVHAASNFKVDKRIEIPDNHQKVC 110
QY 7 ---CPPCAPPELLGGPSVFLPPPKDMLSRTPETVCVVVDVSHEDPEVKFNWYVDGV 63
DB 111 MSKCPKCAPPELLGGPSVFLPPPKDMLSRTPETVCVVVDVSHEDPEVKFNWYMDGV 170
QY 64 EVHNATKPREEQNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIETKISRAKG 123
DB 171 EVRTATPREEQNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIETKISRAKG 230
QY 124 PREPQVTLPPSRDELTKNOVSTCLVKGYPSDIAVWESNGQP--ENNYKTTPPVLD 181
DB 231 SQEPQVVLAPHDELKSKSVTCLVKDFYPEINIEWQSGOPELETYKYSTTQAQDS 290
QY 182 DGSFELYSKLTVDKSRWQGNVFSVMHEALHNHYTKQSLSPGK 228
DB 291 DGSFELYSKLTVDKSRWQGNVFSVMHEALHNHYTKQSLSPGK 337

RESULT 5
Q8R3V9 PRELIMINARY; PRT; 469 AA.
AC Q8R3V9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 52.0 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024405; AAH24405.1; -
DR HSSP; P01842; 7FAB.
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003600; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig_4.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGcl; 2.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 469 AA; 51976 MW; 534793F155D05457 CRC64;

Query Match      62.7%; Score 837; DB 11; Length 469;
Best Local Similarity 52.3%; Pred. No. 2.8e-25;
Matches 145; Conservative 41; Mismatches 35; Indels 56; Gaps 3;

QY 5 HT-----CPP--CP 11
DB 196 HTFPAVLQSDLYTLSSSVTPSPSTFQVTCNVHPASSTKVDKIVPRDCGCKPCICT 255
QY 12 APPELLGGPSVFLPPPKDMLSRTPETVCVVVDVSHEDPEVKFNWYVDGVVHNAKT 71
DB 256 VPEV---SSVFIPPPKDVLTITLTPKVCVVVDISKDDPEVQFSWFVDVHVHTAQK 312
QY 72 PREEQNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIETKISKAGQPREQVYT 131
DB 313 PREEQNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIETKISKAGQPREQVYT 372
QY 132 LPPSRDELTKNOVSTCLVKGYPSDIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKL 191
DB 373 IPPPKQMAKDKVSLTCLMTDFFPEDITVWQWNGQPAENYKNTQPIMDTDGSFYISK 432
QY 192 TVDKSRWQGNVFSVMHEALHNHYTKQSLSPGK 228
DB 433 NVOKSNWEAGNTFTCSVLHGLHNHHTKSLSHSPGK 469

Query Match      62.7%; Score 836; DB 11; Length 463;
Best Local Similarity 52.0%; Pred. No. 3e-25;
Matches 144; Conservative 42; Mismatches 35; Indels 56; Gaps 3;

QY 5 HT-----CPP--CP 11
DB 190 HTFPAVLQSDLYTLSSSVTPSPSTFQVTCNVHPASSTKVDKIVPRDCGCKPCICT 249
QY 12 APPELLGGPSVFLPPPKDMLSRTPETVCVVVDVSHEDPEVKFNWYVDGVVHNAKT 71
DB 250 VPEV---SSVFIPPPKDVLTITLTPKVCVVVDISKDDPEVQFSWFVDVHVHTAQ 306
QY 72 PREEQNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIETKISKAGQPREQVYT 131
DB 307 PREEQNSTYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIETKISKAGQPREQVYT 366
QY 132 LPPSRDELTKNOVSTCLVKGYPSDIAVWESNGQPENNYKTTPPVLDSDGSFFLYSKL 191
DB 367 IPPPKQMAKDKVSLTCLMTDFFPEDITVWQWNGQPAENYKNTQPIMDTDGSFYISK 426
QY 192 TVDKSRWQGNVFSVMHEALHNHYTKQSLSPGK 228
DB 427 NVOKSNWEAGNTFTCSVLHGLHNHHTKSLSHSPGK 463

RESULT 7
Q8R1A4 PRELIMINARY; PRT; 437 AA.
AC Q8R1A4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Gamnal heavy chain of Mab7 (Fragment).
GN IGH-4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilde K.G., Yu X., Ekramoddoullah A.K.M., Misra S.;
RT "Cloning of cDNAs encoding for anti-white pine blister rust monoclonal
```

```

Db 425 DGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNNHYTKQSLSPGK 471
RESULT 2
Q96PQ8 PRELIMINARY; PRT; 701 AA.
AC Q96PQ8
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Factor VII active site mutant immunoconjugate.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21477448; PubMed=11593034;
RA Hu Z., Garen A.;
RT "Targeting tissue factor on tumor vascular endothelial cells and tumor
RT cells for immunotherapy in mouse models of prostatic cancer.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12180-12185(2001).
DR EMBL: AF272774; AAK58686.1; -.
DR InterPro: IPR000152; ASX_Hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001254; Ser_protease_Try.
DR InterPro: IPR000294; VitK_dep_GLA.
DR Pfam: PF00008; EGF_2.
DR Pfam: PF00594; gla; 1.
DR Pfam: PF00047; Ig; 2.
DR Pfam: PF00089; trypsin; 1.
DR SMART: SM00181; EGF_2.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
DR PROSITE: PS01187; EGF_Ca; UNKNOWN_1.
DR PROSITE: PS00011; GLU CARBOXYLATION; UNKNOWN_1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE: PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 701 AA; 77826 MW; 94AC6CEB42CC992F CRC64;

Query Match 92.4%; Score 1233; DB 4; Length 701;
Best Local Similarity 100.0%; Pred. No. 5.7e-41;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DKTHTCPCPAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61
Db 475 DKTHTCPCPAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 534
QY 62 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVKSNKALPAPIETISKAK 121
Db 535 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVKSNKALPAPIETISKAK 594
QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 181
Db 595 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPPVLD 654
QY 182 DGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNNHYTKQSLSPGK 228
Db 655 DGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNNHYTKQSLSPGK 701

RESULT 3
Q8TC63 PRELIMINARY; PRT; 473 AA.
AC Q8TC63
DT 01-JUN-2002 (TReMBLrel. 21, Created)

```

```

DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hypothetical 52.0 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98383416; PubMed=9717671;
RA Wagner B., Overesch G., Sheoran A., Holmes M., Richards C.,
RA Leibold W., Radbruch A.;
RT "Organization of the equine immunoglobulin heavy chain constant region
RT genes. III. Alignment of c-mu, c-gamma, c-epsilon and c-alpha genes.";
RL Immunobiology 199:105-119(1998).
DR EMBL: AJ300675; CAC44624.1; -.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig; 2.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_2.
FT NON_TER 1
SQ SEQUENCE 337 AA; 37438 MW; A60BF2B01DEFD1F6 CRC64;

Query Match 85.7%; Score 1143; DB 4; Length 473;
Best Local Similarity 84.7%; Pred. No. 7.8e-38;
Matches 211; Conservative 9; Mismatches 6; Indels 23; Gaps 1;

QY 3 KTHT-----CPCPAPPELLGGPSVFLFPPPKDTLMISRTPE 39
Db 225 KTYCNVDHKKPSNTKVDKRVESKYGPPCPAPPELLGGPSVFLFPPPKDTLMISRTPE 284
QY 40 VTCVVDVSHEDPEVKFNWYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKE 99
Db 285 VTCVVDVSHEDPEVKFNWYVDGVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKE 344
QY 100 YKCKVSNKALPAPIETISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIA 159
Db 345 YKCKVSNKGLPSSIEKTIISRAKGPQREPVYITLPPSQEEMTKNQVSLTCLVKGFYPSDIA 404
QY 160 VEWESNGQPENNYKTPPPVLDSDGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNNHYTQ 219
Db 405 VEWESNGQPENNYKTPPPVLDSDGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNNHYTQ 464
QY 220 KSLSLSPGK 228
Db 465 KSLSLSLGK 473

RESULT 4
Q95M34 PRELIMINARY; PRT; 337 AA.
AC Q95M34
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Immunoglobulin gamma 1 heavy chain constant region
DE (Fragment).
GN IGHC1.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98383416; PubMed=9717671;
RA Wagner B.;
RA Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98383416; PubMed=9717671;
RA Wagner B., Overesch G., Sheoran A., Holmes M., Richards C.,
RA Leibold W., Radbruch A.;
RT "Organization of the equine immunoglobulin heavy chain constant region
RT genes. III. Alignment of c-mu, c-gamma, c-epsilon and c-alpha genes.";
RL Immunobiology 199:105-119(1998).
DR EMBL: AJ300675; CAC44624.1; -.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig; 2.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_2.
FT NON_TER 1
SQ SEQUENCE 337 AA; 37438 MW; A60BF2B01DEFD1F6 CRC64;

```

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:21 ; Search time 36.1765 Seconds
(without alignments)
1401.120 Million cell updates/sec

Title: 2LINK7
Perfect score: 1334
Sequence: 1 MDRHTCPCPCAPPELLGGPS.....GKDWLKAFAVDKVAEKLKEAF 246

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.0

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1233	92.4	471	4	Q8TC77
2	1233	92.4	701	4	Q96PQ8
3	1143	85.7	473	4	Q8TC63
4	914	68.5	337	6	Q95M34
5	837	62.7	469	11	Q8R3V9
6	836	62.7	463	11	Q9RLC4
7	831	62.3	437	11	Q9RLA4
8	816	61.2	473	11	Q9DLA4
9	811	60.8	473	11	Q99L25
10	810	60.7	468	11	Q99L31
11	788	59.1	474	11	Q8R3H6
12	787	59.0	473	11	Q91Z05
13	486	36.4	34350	4	Q8WZ42
14	466	34.9	26926	4	Q10466
15	466	34.9	26926	4	Q8WZB3
16	417.	31.3	17352	5	Q95YM2

17	405	30.4	16215	5	Q9NFS3	Q9nfs3 drosophila
18	388	29.1	597	4	Q9BQB8	Q9bqb8 homo sapien
19	388	29.1	597	4	Q9BUI0	Q9bui0 homo sapien
20	382	28.6	597	4	Q96BB9	Q96bb9 homo sapien
21	382	28.6	618	4	Q96AA6	Q96aa6 homo sapien
22	379	28.4	588	4	Q8WDX4	Q8wux4 homo sapien
23	378	28.3	7962	4	Q10465	Q10465 homo sapien
24	377	28.3	614	4	Q96GA6	Q96ga6 homo sapien
25	376	28.2	613	4	Q96EV0	Q96ev0 homo sapien
26	376	28.2	613	4	Q8WUK1	Q8wuk1 homo sapien
27	372	27.9	375	4	Q9BSZ1	Q9bsz1 homo sapien
28	364	27.3	613	11	Q8VCX7	Q8vcx7 mus musculu
29	363	27.2	6831	5	Q23550	Q23550 caenorhabdi
30	363	27.2	7160	5	Q23551	Q23551 caenorhabdi
31	362	27.1	15281	3	Q09164	Q09164 tolypocladi
32	356	26.7	4824	5	Q95YM1	Q95ym1 procambarus
33	356	26.7	7107	5	Q9V4F7	Q9v4f7 drosophila
34	352	26.5	5636	4	Q96RW7	Q96rw7 homo sapien
35	352	26.4	6658	5	Q76281	Q76281 drosophila
36	349	26.2	4796	5	Q9NL88	Q9nl88 drosophila
37	348	26.1	6632	5	Q01761	Q01761 caenorhabdi
38	348	26.1	8563	2	Q54297	Q54297 streptomyce
39	348	26.1	13055	5	Q09165	Q09165 caenorhabdi
40	347	26.0	6632	5	Q17362	Q17362 caenorhabdi
41	346	25.9	4796	5	Q9W055	Q9w055 drosophila
42	346	25.9	5604	4	Q8WZ53	Q8wz53 homo sapien
43	344	25.8	5198	5	Q76518	Q76518 caenorhabdi
44	338	25.3	4816	5	Q8T103	Q8t103 bombyx mori
45	337	25.3	9376	2	O85168	O85168 pseudomonas

ALIGNMENTS

RESULT 1

Q8TC77	PRELIMINARY;	PRT;	471 AA.
ID	Q8TC77	PRELIMINARY;	PRT;
AC	Q8TC77;		
DT	01-JUN-2002 (Tremblrel. 21, Created)		
DT	01-JUN-2002 (Tremblrel. 21, Last sequence update)		
DE	Hypothetical 51.8 kDa protein.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=SPLEEN;		
RA	Strausberg R;		
RL	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC024289; AAH24289.1;		
KW	Hypothetical protein.		
SQ	SEQUENCE 471 AA; 51791 MW; 388F7F4CF588660E CRC64;		
Query Match	92.4%;	Score 1233;	DB 4; Length 471;
Best Local Similarity	100.0%;	Pred. No. 1.5e-41;	
Matches 227;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	2 DKHTCPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61		
Db	245 DKHTCPCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 304		
QY	62 GVEVHNAKTRPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121		
Db	305 GVEVHNAKTRPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 364		
QY	122 GQREPOVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPPVLD 181		
Db	365 GQREPOVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGQPENNYKTPPPVLD 424		
QY	182 DGSFFFLYSKLTVDKSRWQQGNVSCFVSHEALHNHYTQKSLSLSPGK 228		

```
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 144 204
FT DISULFID 250 308
FT MOD_RES 330 330 REMOVED POST-TRANSLATIONALLY.
SQ SEQUENCE 330 AA; 36389 MW; B84361C5445A8684 CRC64;

Query Match 60.9%; Score 813; DB 1; Length 330;
Best Local Similarity 52.9%; Pred. No. 2.4e-20;
Matches 148; Conservative 30; Mismatches 46; Indels 56; Gaps 2;

Qy 5 HT-----CPP- 9
Db 51 HTFPAVLOSDLYTLSSSVTVTSSTWPSQSIITCNVAHPASSTKVDKKIEPRGPTIKPCPPC 110
Qy 10 -CPAPELLGGPSVFLLPPKPKDLMISRTPEVTCVVVDVSHEDPEVKENWYVDGVEVHNA 68
Db 111 KCPANLLGGPSVFIFPKIKDVLMSLSPIVTCVVVDVSEDDPDVQISWFVNNVEVHTA 170
Qy 69 KTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQ 128
Db 171 QTQTHREDYNSTLRVVSALPIQHODWMSGKEFKCKVNNKDLPAPIERTISKPKGSRAPQ 230
Qy 129 VYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFELY 188
Db 231 VVLPPEPEEETKQVTLTCMTVDFMPEDIYVEWTNNGKTELNYKNTPEVLDSDGSYFMY 290
Qy 189 SKLTVDKSRWQGNVFCGVMHEALHNHYTKLSLSPGK 228
Db 291 SKLRVEKKNWVERNSYSCSVVHEGLNHHHTTKSFSRTFGK 330
```

Search completed: April 21, 2003, 10:43:10
Job time : 10.4059 secs

DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig gamma-2C chain C region.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=88166903; PubMed=3127222;
 RX Brueggemann M., Delmastro-Galfré P., Waldmann H., Calabi F.;
 RT "Sequence of a rat immunoglobulin gamma 2c heavy chain constant
 region cDNA: extensive homology to mouse gamma 3.";
 RL Eur. J. Immunol. 18:317-319(1988).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X07189; CAA30169.1; -
 DR PIR; S00847; S00847.
 DR HSSP; P01842; 7FAB.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003597; Ig_cl.
 DR InterPro; IPR003600; Ig_like.
 DR Pfam; PF00047; Ig; 2.
 DR SMART; SM00410; Ig_like; 1.
 DR SMART; SM00407; IGcl; 2.
 DR PROSITE; PS00290; IG_MHC; 1.
 DR Immunoglobulin domain; Immunoglobulin C region.
 FT NON_TER 1 1
 FT DOMAIN 1 97 CH1.
 FT DOMAIN 98 113 HINGE.
 FT DOMAIN 114 222 CH2.
 FT DOMAIN 223 329 CH3.
 FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
 FT DISULFID 27 82
 FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 113 113 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 143 203
 FT DISULFID 249 307
 SQ SEQUENCE 329 AA; 36571 MW; 5FCD7B7933850773 CRC64;
 Query Match 61.3%; Score 818; DB 1; Length 329;
 Best Local Similarity 57.9%; Pred. No. 1.6e-20;
 Matches 146; Conservative 42; Mismatches 38; Indels 26; Gaps 2;
 QY 3 KTHTC-----PP---CPAPELLGGPSVFLPFPKPKDTLMISR 36
 DB 78 QTVCSVAHPATKSNLIKRIEPRRPKRPTDTCSDDLGRSVFIFFPKPKDILMITL 137
 QY 37 TPVTCVVVDVSHEDPEVFANFVVDGVEVHNAKTKPREQYNSTYRVSWLTVLHQQDWLN 96
 DB 138 TPVTCVVVDVSEEDPQSFWEVDNVRVFTAGTQPHHEQLNGTFRVSTLHQQDWMS 197
 QY 97 GKYYKCKVSKALPAPTEIKTSKAKGPQVYTPPPSRDELTKNOVSLTCLVKGFPYS 156
 DB 198 GKFEKCKVANKDLPSPIETKISKPRGKARTPQVYTPPPPREQMSKKNKSLTCLVTSFYPA 257
 QY 157 DIAVENESGCPENNVKTPPVLDSDGSFPLYSLKLVKDSRWQOGNVFSCVWHEALHNN 216
 DB 258 SISVERNGEGLFQDYKNTLPVLDSDGSFPLYSLKLVKDSRWQOGNVFSCVWHEALHNN 317
 QY 217 YTKSLSLSPGK 228
 DB 318 HTQKNLSRSPGK 329
 RESULT 15

GCAN_MOUSE
 ID GCAN_MOUSE STANDARD; PRT; 330 AA.
 AC P01863;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Ig gamma-2A chain C region, A allele.
 DE Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=81076554; PubMed=6777755;
 RX Sikorav J.-L., Auffray C., Rougeon F.;
 RT "Structure of the constant and 3' untranslated regions of the murine
 RT Balb/c gamma 2a heavy chain messenger RNA.";
 RL Nucleic Acids Res. 8:3143-3155(1980).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=81198976; PubMed=6262729;
 RA Yamawaki-Kataoka Y., Miyata T., Honjo T.;
 RT "The complete nucleotide sequence of mouse immunoglobulin gamma 2a gene
 RT and evolution of heavy chain genes: further evidence for intervening
 RT sequence-mediated domain transfer.";
 RL Nucleic Acids Res. 9:1365-1381(1981).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=81223894; PubMed=6787604;
 RA Ollo R., Auffray C., Morchamps C., Rougeon F.;
 RT "Comparison of mouse immunoglobulin gamma 2a and gamma 2b chain genes
 RT suggests that exons can be exchanged between genes in a multigenic
 RT family.";
 RL Proc. Natl. Acad. Sci. U.S.A. 78:2442-2446(1981).
 RN [4]
 RP MYELOMA PROTEIN MOPC 173.
 RX MEDLINE=74175517; PubMed=4831970;
 RA Bourgois A., Fougereau M., Rocca-Serra J.;
 RT "Determination of the primary structure of a mouse IgG2a
 RT immunoglobulin: amino-acid sequence of the Fc fragment. Implications
 RT for the evolution of immunoglobulin structure and function.";
 RL Eur. J. Biochem. 43:423-435(1974).
 RN [5]
 RP DISULFIDE BONDS.
 RX MEDLINE=73056887; PubMed=4565406;
 RA de Preval C., Fougereau M.;
 RT "Determination of the primary structure of a mouse gamma G2a
 RT immunoglobulin. Identification of the disulfide bridges.";
 RL Eur. J. Biochem. 30:452-462(1972).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; V00798; CAA24178.1; -
 DR PIR; A02152; G2MSA.
 DR HSSP; P01842; 7FAB.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003597; Ig_cl.
 DR InterPro; IPR003600; Ig_like.
 DR Pfam; PF00047; Ig; 2.
 DR SMART; SM00410; Ig_like; 1.
 DR SMART; SM00407; IGcl; 2.
 DR PROSITE; PS00290; IG_MHC; 1.
 DR Immunoglobulin domain; Immunoglobulin C region.
 KW NON_TER 1 1
 FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
 FT DISULFID 27 82
 FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 107 107


```

DR EMBL; V00793; CAA24174.1;
DR PIR; B02159; G1NSM.
DR HSP; P01842; 7FAB.
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_c1.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; Igc1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing; Transmembrane.
FT NON_TER 1
FT DOMAIN 1 97
FT DOMAIN 98 110
FT DOMAIN 111 217
FT DOMAIN 218 324
FT DISULFID 27 82
FT DISULFID 102 102
FT DISULFID 104 104
FT DISULFID 107 107
FT DISULFID 109 109
FT DISULFID 138 198
FT CARBOHYD 174 174
FT DISULFID 244 302
FT TRANSMEM 340 357
FT DOMAIN 338 393
SQ SEQUENCE 393 AA; 43386 MW; 4CC88343B7A1CE27 CRC64;

Query Match 61.7%; Score 823; DB 1; Length 393;
Best Local Similarity 51.8%; Pred. No. 1.9e-20;
Matches 145; Conservative 47; Mismatches 42; Indels 46; Gaps 5;

QY 7 CPP--CPAPELGGPSVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVE 64
DB 104 CKPCICTVEV--SSVFIFPKPKDVLITLTPKVTVCVVVDISKDDPEVQSFVWVDVE 160
QY 65 VHNATKPREEQNSTYRVSVTLVHQLDNLNGKEYCKVSNKALPAPIEKTIISKAKGP 124
DB 161 VHTAQTPREEQNSTFRSVSELPIMHQLDNLNGKEFKCRVNSAFAFPAPIEKTIISKGR 220
QY 125 REPQVYTLPPSRDELTKNOVSLTCLVKGYPSDIAVEWESNGPENNYKTTTPVLDSDGS 184
DB 221 KAPOVYTIPTPEQAKDKVSLTCHITDFPEDIYVQWQNGQPAENYKNTQPIININGS 280
QY 185 FFLYSLKLTVDKSRWQGNVFCSSVMHEALHNHYTKSLSPGKD----- 229
DB 281 YFYVSKLVNOKSWNEAGNTFTCSVLHGLNHHHTKSLSPGQLDTCACAEQDGLDG 340
QY 230 -W-----LKFYDKVAEKLKE 244
DB 341 LMTTITIFISLLSVCYSAATLFPKRVWIFSSVVE-LKQ 379

RESULT 13
GCAM_MOUSE
ID GCAM_MOUSE STANDARD; PRT; 399 AA.
AC P01865;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2A chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82222190; PubMed=6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
RT immunoglobulin gamma chains."
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
CC -|- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA

```

```

CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
CC BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
CC IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
CC SEGMENT OF MU CHAINS.
CC -|- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 1-329 IS ASSUMED TO BE
CC IDENTICAL WITH THE CORRESPONDING REGION OF THE SECRETED FORM OF
CC THE A ALLELE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J00471; AAB59661.1; ALT_INIT.
CC PIR; A02154; G2MSAM.
CC HSP; P01842; 7FAB.
CC MGD; MGI:96443; Igh-1.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR003597; Ig_c1.
CC InterPro; IPR003600; Ig_like.
CC Pfam; PF00047; Ig; 2.
CC SMART; SM00410; IG_like; 1.
CC SMART; SM00407; Igc1; 2.
CC PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1
FT DISULFID 15 15
FT DISULFID 27 82
FT DISULFID 107 107
FT DISULFID 110 110
FT DISULFID 112 112
FT DISULFID 144 204
FT DISULFID 250 308
FT TRANSMEM 346 363
FT DOMAIN 364 399
FT CARBOHYD 180 180
SQ SEQUENCE 399 AA; 44020 MW; 4C38138BFAED3FF0 CRC64;

Query Match 61.4%; Score 819; DB 1; Length 399;
Best Local Similarity 54.3%; Pred. No. 2.7e-20;
Matches 152; Conservative 33; Mismatches 52; Indels 43; Gaps 4;

QY 7 CPP--CPAPELGGPSVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVE 64
DB 107 CPPCKPAPNLLGGPSVFIFFPKIKDVLMSISPIVTCVVVDVSEDDPQVQISWFVNNVE 166
QY 65 VHNATKPREEQNSTYRVSVTLVHQLDNLNGKEYCKVSNKALPAPIEKTIISKAKGP 124
DB 167 VHTAQTPREEQNSTLRVVSALPIQHDQMSGKEFKCRVNSAFAFPAPIERISPKGSV 226
QY 125 REPQVYTLPPSRDELTKNOVSLTCLVKGYPSDIAVEWESNGPENNYKTTTPVLDSDGS 184
DB 227 RAPQVYVLPPEEEMTKQVITLCVMTDFMPEDIYVEWTNNGKTELNYKNTPEVLDSDGS 286
QY 185 FFLYSLKLTVDKSRWQGNVFCSSVMHEALHNHYTKSLSPGKD----- 229
DB 287 YFYVSKLVNOKSWNEAGNTFTCSVLHGLNHHHTKSLSPGQLDTCACAEQDGLDG 346
QY 230 -W-----LKFYDKVAEKLKE 244
DB 347 LMTTITIFISLLSVCYSAATLFPKRVWIFSSVVE-LKQ 385

RESULT 14
GCC_RAT
ID GCC_RAT STANDARD; PRT; 329 AA.
AC P20762;
DT 01-FEB-1991 (Rel. 17, Created)

```

RA Rogers J., Clarke P., Salsner W.;
 RT "Sequence analysis of cloned cDNA encoding part of an immunoglobulin
 RL heavy chain."; Res. 6:3305-3321(1979).
 RN [4]
 RP SEQUENCE (MYELOMA PROTEIN MOPC 21).
 RX MEDLINE=78242288; PubMed=98524;
 RA Adetugbo K.;
 RT "Evolution of immunoglobulin subclasses. Primary structure of a
 RL murine myeloma gamma1 chain."; J. Biol. Chem. 253:6068-6075(1978).
 RN [5]
 RP DISULFIDE BONDS (MOPC 21).
 RX MEDLINE=73008889; PubMed=5073237;
 RA Svastil J., Milstein C.;
 RT "The disulphide bridges of a mouse immunoglobulin G1 protein.";
 RL Biochem. J. 126:837-850(1972).
 RN [6]
 RP This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; V00793; CAA24172.1; -;
 DR EMBL; V00793; CAA24173.1; -;
 DR EMBL; V00793; CAA24174.1; -;
 DR EMBL; V00793; CAA24175.1; -;
 DR EMBL; V00795; CAA24176.1; -;
 DR PIR; A02159; GIMS.
 DR HSSP; P01842; 7FAB.
 DR GlycoSuiteDB; P01868; -;
 DR MGD; MGI:96446; Igh-4.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003597; Ig-cl.
 DR Pfam; PF00047; Ig; 3.
 DR SMART; SM00407; Igcl; 2.
 DR PROSITE; PS00290; IG_MHC; 1.
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
 KW Alternative splicing.
 FT NON_TER 1 1
 FT DOMAIN 1 97 CH1.
 FT DOMAIN 98 110 HINGE.
 FT DOMAIN 111 217 CH2.
 FT DOMAIN 218 324 CH3.
 FT DISULFID 27 82
 FT DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).
 FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 138 198
 FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .).
 FT FTID=CAR_000055.
 FT DISULFID 244 302
 FT MOD_RES 324 324 REMOVED POST-TRANSLATIONALLY.
 FT CONFLICT 276 276 N -> D (IN REF. 3).
 FT CONFLICT 278 278 N -> D (IN REF. 3).
 SQ SEQUENCE 324 AA; 35704 MW; A338812F3D1F2C93 CRC64;
 Query Match 61.7%; Score 823; DB 1; Length 324;
 Best Local Similarity 51.3%; Pred. No. 1e-20;
 Matches 142; Conservative 44; Mismatches 35; Indels 56; Gaps 3;
 5 HT-----
 11
 51 HTFAVLQSDLYLTSSSVTPSPRSETVTCNVAPASSTKVDDKIVPRDCGCKPCICT 110
 12 APELLGSPSVFLPPRPKDTLMISRPETVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTK 71
 111 VPEV---SSVFIIPPKPKVDLTITLTPKVTCCVVVDISKDDPEQFSWFDVDEVHTAQTK 167

QY 72 PREEQYNSTYRVSVLTVLHODWLNKGKCKYKSNKALPAPIEKTISKARGQPREPOVYT 131
 DB 168 PREEQFNSTFRSVSELPIMHODWLNKGKFKCRVNSAFAFPAPIEKTISKGRKAPOVYT 227
 QY 132 LPFSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSGFFLYSKL 191
 DB 228 IPPPEQMAKDKVSLTCMTITDFPEDITVEMQWNGQPAENYKNTQPIMTNGSVFYVSKL 287
 QY 192 TVDKSRWQQGVNFCSCVMHEALHNHYTKQSLSPGK 228
 DB 288 NVQKSNEAGNFTCSVLHGLHNHHTKSLSHSPGK 324
 RESULT 12
 GCIM_MOUSE STANDARD; PRT; 393 AA.
 AC P01869;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Ig gamma-1 chain C region, membrane-bound form.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=80045036; PubMed=115593;
 RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
 RA Takahashi N., Mano Y.;
 RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
 RL gamma 1 chain gene";
 RL Cell 18:559-568(1979).
 RN [2]
 RP SEQUENCE OF 323-393 FROM N.A.
 RX MEDLINE=82197626; PubMed=6804950;
 RA Tyler B.M., Cowman A.F., Gerondakis S.D., Adams J.M., Bernard O.;
 RT "mRNA for surface immunoglobulin gamma chains encodes a highly
 RT conserved transmembrane sequence and a 28-residue intracellular
 RT domain";
 RN Proc. Natl. Acad. Sci. U.S.A. 79:2008-2012(1982).
 RN [3]
 RP SEQUENCE OF 323-366 FROM N.A.
 RX MEDLINE=82115295; PubMed=6799207;
 RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,
 RA Eisenberg D., Wall R.;
 RT "Gene segments encoding transmembrane carboxyl termini of
 RT immunoglobulin gamma chains";
 RN Cell 26:19-27(1981).
 RN [4]
 RP SEQUENCE OF 1-44 FROM N.A.
 RX MEDLINE=82222190; PubMed=6283537;
 RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
 RT "Nucleotide sequences of gene segments encoding membrane domains of
 RT immunoglobulin gamma chains";
 RN Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
 CC -!- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA
 CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
 CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
 CC BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
 CC IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
 CC SEGMENT OF MU CHAINS
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; V00793; CAA24172.1; -;
 DR EMBL; V00793; CAA24173.1; -;

DR EMBL; V01526; CAA24767.1; ALT_SEQ.
 DR PIR; A02155; G3MSM.
 DR HSSP; P01857; IFC1.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003597; Ig-cl.
 DR InterPro; IPR003600; Ig-like.
 DR Pfam; PF00047; Ig; 3.
 DR SMART; SM00410; Ig-like; 1.
 DR SMART; SM00407; IGcl; 2.
 DR PROSITE; PS00290; IG_MHC; 1.
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
 Transmembrane; Alternative splicing.
 FT NON_TER 1 1
 FT DOMAIN 1 97 CH1.
 FT DOMAIN 98 113 HINGE.
 FT DOMAIN 114 223 CH2.
 FT DOMAIN 224 327 CH3.
 FT TRANSMEM 346 362 POTENTIAL.
 FT DOMAIN 363 398 CYTOPLASMIC (POTENTIAL).
 FT CONFLICT 333 333 E -> G (IN REF. 2).
 FT CONFLICT 342 342 E -> Q (IN REF. 2).
 FT CONFLICT 388 388 P -> F (IN REF. 2).
 SQ SEQUENCE 398 AA; 43929 MW; CF7F264B50A41B95 CRC64;

Query Match 62.8%; Score 838; DB 1; Length 398;
 Best Local Similarity 50.5%; Pred. No. 6e-21;
 Matches 151; Conservative 37; Mismatches 41; Indels 70; Gaps 3;
 3 KTHCTC-----PP---CPAPPELLGGPSVFLPPKPKDTLMIS 35
 77 QTVCINVAHPASKTELKRIEPRIPKPTPGSSCPPNIGGPSVFPPKPKDLMIS 136
 36 RPEVTCVVDVSHEDPEVKWYVDGVEVHNAKTKPREEOYNSTRYVSVLTVLHODWL 95
 137 LTPKVTVCVVDVSEDDPDVHVSFWVDNKEVHTAWTPREAQYNSTRFVSALPIQHQQDM 196
 96 NGKEYCKVKSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNOVSLTCLVKGY 155
 197 RGKEFKCKVKNKALPAPIERTISKPKRAQTPQVYTIPTPREQMSKKVKVSLTCLVNTFS 256
 156 SDIAVEWESNGOPENNYKTPPVLDSGFFLYSKLTVDKSRWQGNVFCSCVMHEALHN 215
 257 EASVEWERNGLGEQDYKNTPPILDSGDTYFLYSLKLTVDTSWLGGEFTCSVMHEALHN 316
 216 HYTKSLSLSPGK-----WL 231
 317 HHTQNLRSRPELNETCAEAQDGLDGLWTITIFISLFLSVCYASVTLFRVKWI 375

RESULT 10
 GCL_RAT
 ID GCL_RAT STANDARD; PRT; 326 AA.
 AC P20759;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig gamma-1 chain C region.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89232738; PubMed=3149946;
 RA Brueggemann M.;
 RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family."
 RL Gene 74:473-482(1989).
 DR PIR; PS0017; PS0017.
 DR HSSP; P01842; 7FAB.
 DR InterPro; IPR003006; Ig_MHC.
 DR Pfam; PF00047; Ig; 3.
 DR SMART; SM00407; IGcl; 2.

DR PROSITE; PS00290; IG_MHC; 1.
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
 FT NON_TER 1 1
 FT DOMAIN 1 97 CH1.
 FT DOMAIN 98 112 HINGE.
 FT DOMAIN 113 219 CH2.
 FT DOMAIN 220 326 CH3.
 FT DISULFID 27 82
 FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 140 200
 FT DISULFID 246 304
 FT CARBOHYD 176 176
 SQ SEQUENCE 326 AA; 35946 MW; 013BAB45EF49B9DA CRC64;
 Query Match 61.8%; Score 825; DB 1; Length 326;
 Best Local Similarity 51.4%; Pred. No. 9.1e-21;
 Matches 144; Conservative 41; Mismatches 35; Indels 60; Gaps 3;
 5 HT-----CPGPCP 11
 51 HTFPAVLQGLYLTSSVTPSSWTSPQTVTCNVAHPASTKVDKKIYPRNCGGDKPC- 109
 12 APELLGG---PSVFLPPKPKDTLMISRTPEVTCVVDVSHEDPEVKENWYVDGVEVHN 68
 110 ---ICTGSEVSVFIFPPKPKDVLITLTPKVTVCVVDVSDQDDPEVHFSWFDVDEVHTA 166
 69 KTKPREEOYNSTRYVSVLTVLHODWLNKKEYCKVKSNKALPAPIEKTISKAKGQPREPQ 128
 167 QTRPPEQFNSTRFSVSELPILHODWLNRTFRCTSAAPSPLEKTSKPEGTQVPH 226
 129 VTLPPSRDELTKNOVSLTCLVKGYPSDIAVEWESNGOPENNYKTPPVLDSGDSFFLY 188
 227 VYTSPTKEMTQNEVSTICWVKGYPPDIYVEWQGNQPNQENYKNTPTDGTGSGFLY 286
 189 SKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSLSPGK 228
 287 SKLNVKKEKQGNFTTCSVLHGLHNHTEKLSHSPGK 326
 RESULT 11
 GCL_MOUSE
 ID GCL_MOUSE STANDARD; PRT; 324 AA.
 AC P01868;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Ig gamma-1 chain C region.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=80045036; PubMed=115593;
 RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
 RA Takahashi N., Mano Y.;
 RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
 gamma 1 chain gene."
 RL Cell 18:559-568(1979).
 RN [2]
 RP SEQUENCE OF 76-324 FROM N.A. (MYELOMA PROTEIN MOPC 31C).
 RX MEDLINE=80202559; PubMed=6769752;
 RA Obata M., Yamawaki-Kataoka Y., Takahashi N., Kataoka T., Shimizu A.,
 RA Mano Y., Seidman J.G., Peterlin B.M., Leder P., Honjo T.;
 RT "Immunoglobulin gamma 1 heavy chain gene: structural gene sequences
 cloned in a bacterial plasmid."
 RL Gene 9:87-97(1980).
 RN [3]
 RP SEQUENCE OF 70-322 FROM N.A. (MYELOMA PROTEIN MOPC 21).
 RX MEDLINE=80012837; PubMed=113776;

```
CC ----- EMBL; J00451; NOT_ANNOTATED_CDS.
DR PIR; B02156; G3MSC.
DR HSSP; P01857; 1FC1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; Ig_like; 1.
DR SMART; SM00407; IGc1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
SQ SEQUENCE 329 AA; 36228 MW; F45827174182BAD6 CRC64;

Query Match 63.6%; Score 848; DB 1; Length 329;
Best Local Similarity 60.1%; Pred. No. 1.5e-21;
Matches 152; Conservative 35; Mismatches 39; Indels 27; Gaps 2;

QY 3 KTHFC-----pp-----CPAPELLGGPSVFLPPPKDGLMIS 35
Db 77 QTICVNAHPASKTELKRIEPIKPPSPGSCPPGNILGSPSVFIPPKPKDALMIS 136
QY 36 RTEPTCVVVDVSHEDPEVKFNMYVDGVEVHNNAKTPREQYNSTYRVVSVLTVLHQDWL 95
Db 137 LTPKVCVVVDVEDDPPDVHVSFVDNKEVHTAQTQPREAQYNSTFRVVSALPIQHDWM 196
QY 96 NGKEYCKVSNKALPAPIEKTISKAKGQPREPOVYITLPPSRDELTKNOVSLTCLVKGFTP 155
Db 197 RGKEFKCKVNNKALPAPIERTISKPKGRAQTPQVYITPPPREQMKKSVLTCLVTNFFS 256
QY 156 SDIATVWESNGPENNYKTPPVLDSDGSGFFLYSKLTVDKSRWQGNVFCSCVMHREALHN 215
Db 257 EALSVWERNGEQDYKNTPTPLDSDGTGYFLYSLKLTVDTSWLOGEITFCVSVHREALHN 316
QY 216 HYTKSLSLSPGK 228
Db 317 HHTQKNLSRSPGK 329

RESULT 8
GCB_RAT STANDARD; PRT; 333 AA.
AC P20761;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2B chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=69232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
DR PIR; PS0018; PS0018.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; Ig_like; 1.
DR SMART; SM00407; IGc1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
```

```
FT NON_TER 1 1 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 15 15
FT DISULFID 27 80
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 147 207
FT DISULFID 253 311
SQ SEQUENCE 333 AA; 55P8B64D48D460A6 CRC64;

Query Match 63.0%; Score 840; DB 1; Length 333;
Best Local Similarity 62.1%; Pred. No. 3e-21;
Matches 151; Conservative 34; Mismatches 43; Indels 15; Gaps 2;

QY 1 MDKT-----HTCP-----PCPAPELLGGPSVFLPPPKDGLMISRTPEVTCVV 45
Db 91 VDKKVERRNGGIGHKCPCTCHCKPVPPELLGGPSVFIFPPKPKDILLISQNAKVTCCVV 150
QY 46 DYSHEDPEVKFNMYVDGVEVHNNAKTPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVS 105
Db 151 DYSEEDPDVQSFVNNVEVHTAQTQPREQYNSTFRVVSALPIQHDWMNSGKEFKCKVN 210
QY 106 NKALPAPIEKTISKAKGQPREPOVYITLPPSRDELTKNOVSLTCLVKGFTVPSDIAVWESN 165
Db 211 NKALPSPIEKTISKPKGLYRKPOVYVGMPTTEQLTQVSLTCLTSGFLPNDIGVETSN 270
QY 166 GOPENNYKTPPVLDSDGSGFFLYSKLTVDKSRWQGNVFCSCVMHREALHNHYTKSLSL 225
Db 271 GHIEKYNKTEPVMDSDGSGFFMYSKLNVRSRWDSPAPFCVSVVHGLHNHHVKEKSTRP 330
QY 226 PGK 228
Db 331 PGK 333

RESULT 9
GCB_MOUSE STANDARD; PRT; 398 AA.
AC P03987;
DT 23-OCT-1986 (Rel. 02, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-3 chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;
RA Wells J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBO J. 3:2041-2046(1984).
RN [2]
RP SEQUENCE OF 328-398 FROM N.A.
RX MEDLINE=84041483; PubMed=6314258;
RA Komaromy M., Clayton L., Rogers J., Robertson S., Kettman J.,
RA Wall R.;
RT "The structure of the mouse immunoglobulin in gamma 3 membrane gene
segment.";
RL Nucleic Acids Res. 11:6775-6785(1983).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00451; AAB59655.1; -.
```

SQ SEQUENCE 323 AA; 35404 MW; 69E8AA118D579A8B CRC64;
 Query Match 69.0%; Score 921; DB 1; Length 323;
 Best Local Similarity 71.1%; Pred. No. 4.7e-24;
 Matches 167; Conservative 29; Mismatches 32; Indels 5; Gaps 2;
 QY 1 MDKT---HTC--PPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVK 55
 DB 91 VDKTVAISTCKSPKCPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVK 150
 QY 56 FNTWVDGVEVHNATKPREQYNSTYRVSVLTFLVHODWLNKREYKCKVSNKALPAPIEK 115
 DB 151 FTWYINNEQVTRAPPLREQNFSTIRVSTPLTHQDWLNKREYKCKVSNKALPAPIEK 210
 QY 116 TISKAKGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTT 175
 DB 211 TISKARQPLEPKVITMGPPREELSSRSVSLTCMNGFYPSDIAVEWESNGKAEADNYKTT 270
 QY 176 PVLDSGDFLYSKLVKRVDSKRMQGNVFCVSMHEALHNHYTKQSLSPGK 228
 DB 271 PAVLDSGDFLYNKLKSVPTSEWQGDVFTCSVMHEALHNHYTKQSLSPGK 323
 RESULT 6
 GC2_CAVPO STANDARD; PRT; 329 AA.
 ID GC2_CAVPO
 AC P01862;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig gamma-2 chain C region.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Mystricognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE OF 1-3.
 RA Trischmann T.M.;
 RL Submitted (APR-1975) to the PIR data bank.
 RN [2]
 RP SEQUENCE OF 4-68.
 RX MEDLINE=71058471; PubMed=5538606;
 RA Birstein B.K., Hussain Q.Z., Cebra J.J.;
 RT "Structure of heavy chain from strain 13 guinea pig
 immunoglobulin-G(2). 3. Amino acid sequence of the region around the
 half-cysteine joining heavy and light chains.";
 RL Biochemistry 10:18-25(1971).
 RN [3]
 RP SEQUENCE OF 69-133 AND 312-329.
 RX MEDLINE=71058486; PubMed=5538616;
 RA Turner K.J., Cebra J.J.;
 RT "Structure of heavy chain from strain 13 guinea pig
 immunoglobulin-G(2). II. Amino acid sequence of the carboxyl-terminal
 and hinge region cyanoogen bromide fragments.";
 RL Biochemistry 10:9-17(1971).
 RN [4]
 RP SEQUENCE OF 134-226.
 RX MEDLINE=75036072; PubMed=4429665;
 RA Tracey D.E., Cebra J.J.;
 RT "Primary structure of the CH2 homology region from guinea pig IgG2
 antibodies.";
 RL Biochemistry 13:4796-4803(1974).
 RN [5]
 RP SEQUENCE OF 227-311.
 RX MEDLINE=75036073; PubMed=4609467;
 RA Trischmann T.M., Cebra J.J.;
 RT "Primary structure of the CH3 homology region from guinea pig IgG2
 antibodies.";
 RL Biochemistry 13:4804-4811(1974).
 RN [6]
 RP DISULFIDE BONDS.
 RX MEDLINE=71058474; PubMed=4922544;
 RA Oliveira B., Lamm M.E.;

RT "Interchain disulfide bridges of guinea pig gamma-2-immunoglobulin.";
 RL Biochemistry 10:26-31(1971).
 CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM POOLED SERUM OF STRAIN
 CC 13 INBRED GUINEA PIGS.
 DR PIR: A02151; G2GP.
 DR HSP: P01842; 7FAB.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003597; Ig-cl.
 DR InterPro: IPR003600; Ig-like.
 DR Pfam: PF00047; Ig; 2.
 DR SMART: SM00410; Ig-like; 1.
 DR SMART: SM00407; IG1; 2.
 DR PROSITE: PS00290; IG_MHC; 1.
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
 FT NON_TER 1
 FT DISULFID 16 16 INTERCHAIN (WITH A LIGHT CHAIN).
 FT DISULFID 28 79
 FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 142 202
 FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .).
 FT DISULFID 248 308
 SQ SEQUENCE 329 AA; 36074 MW; 5D231B7164D1FBA9 CRC64;
 Query Match 67.0%; Score 894; DB 1; Length 329;
 Best Local Similarity 68.8%; Pred. No. 4.2e-23;
 Matches 165; Conservative 25; Mismatches 37; Indels 13; Gaps 2;
 QY 1 MDKT-----HTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSH 49
 DB 90 VDKTVPRTZPBPCPCPCPPENLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSH 149
 QY 50 EDPEKFNWYVDGVEVHNATKPREQYNSTYRVSVLTFLVHODWLNKREYKCKVSNKAL 109
 DB 150 DEPEQVTFVDNKPVGNAETKPREQYNSTYRVSVLTFLVHODWLNKREYKCKVSNKAL 209
 QY 110 PAPIETISKAKQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQP- 168
 DB 210 PAPIETISKAKQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQP- 269
 QY 169 -ENNYKTTTPVLDSDGDFLYSKLVKRVDSKRMQGNVFCVSMHEALHNHYTKQSLSPG 227
 DB 270 SEKEYNTPTIEDADGDFLYSKLVKRVDSKRMQGNVFCVSMHEALHNHYTKQSLSPG 329
 RESULT 7
 GC3_MOUSE STANDARD; PRT; 329 AA.
 ID GC3_MOUSE
 AC P22436;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ig gamma-3 chain C region, secreted form.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85027161; PubMed=6092053;
 RA Wells J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
 RA Tucker P.W., Blattner F.R.;
 RT "Structural analysis of the murine IgG3 constant region gene.";
 RL EMBO J. 3:2041-2046(1984).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

RA Pink J.R.L., Buttery S.H., de Vries G.M., Milstein C.;
 RT "Human immunoglobulin subclasses. Partial amino acid sequence of the
 RL Biochem. J. 117:33-47(1970).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: K01316; AAB59394.1; ALT_INIT.
 DR PIR: A02150; G4HU.
 DR HSSP: P01842; 7FAB.
 DR Genew: HGNC:5528; IGHG4.
 DR MIM: 147130;
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003597; Ig_c1.
 DR InterPro: IPR003600; Ig_like.
 DR Pfam: PF00047; Ig; 3.
 DR SMART: SM00410; Ig_like; 1.
 DR SMART: SM00407; Ig_c1; 2.
 DR PROSITE: PS00290; IG_MHC; 2.
 DR Immunoglobulin domain; Immunoglobulin C region.
 KW NON_TER 1 1
 FT DOMAIN 1 98 CH1.
 FT DOMAIN 99 110 HINGE.
 FT DOMAIN 111 220 CH2.
 FT DOMAIN 221 327 CH3.
 FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).
 FT DISULFID 27 83 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
 FT DISULFID 141 201
 FT DISULFID 247 305
 SQ SEQUENCE 327 AA; 35940 MW; 3EBDB811EF208E7A CRC64;
 Query Match 85.6%; Score 1142; DB 1; Length 327;
 Best Local Similarity 84.7%; Pred. No. 1.4e-31;
 Matches 211; Conservative 9; Mismatches 6; Indels 23; Gaps 1;
 QY 3 KHT-----CPPAPELLGGPSVFLPPPKDKTILMISRTPE 39
 Db 79 KTYCNVDVSHEDPEVKFNKYVDGVEVHNKTKPREQYNSTYRVSVLTVLHODWLNKRE 99
 QY 40 VTCVVVDVSHEDPEVKFNKYVDGVEVHNKTKPREQYNSTYRVSVLTVLHODWLNKRE 99
 Db 139 VTCVVVDVSHEDPEVKFNKYVDGVEVHNKTKPREQYNSTYRVSVLTVLHODWLNKRE 198
 QY 100 YKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIA 159
 Db 199 YKCKVSNKGLPSSIEKTISKAKGQPREPQVYTLPPSQEEMTKNQVSLTCLVKGFYPSDIA 258
 QY 160 VEVESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRNQQGNVFSCSVMHEALHNHYTQ 219
 Db 259 VEVESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRNQQGNVFSCSVMHEALHNHYTQ 318
 QY 220 KSLSLSPGK 228
 Db 319 KSLSLSLGK 327
 RESULT 4
 GC3_HUMAN
 ID GC3_HUMAN STANDARD; PRT; 290 AA.
 AC P01860;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DT Ig gamma-3 chain C region (Heavy chain disease protein) (HDC).
 GN IGHG3.

OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE (DISEASE PROTEIN WIS).
 RP MEDLINE=81021548; PubMed=6774747;
 RA Frangione B., Rosenwasser E., Prelli F., Franklin E.C.;
 RT "Primary structure of human gamma 3 immunoglobulin deletion mutant:
 RL gamma 3 heavy-chain disease protein wis.";
 RL Biochemistry 19:4304-4308(1980).
 RN [2]
 RP REVISIONS TO 12-97 (PROTEIN WIS).
 RP MEDLINE=77118561; PubMed=402363;
 RA Michaelson T.E., Frangione B., Franklin E.C.;
 RT "Primary structure of the 'hinge' region of human IgG3. Probable
 RT quadruplication of a 15-amino acid residue basic unit.";
 RL J. Biol. Chem. 252:883-889(1977).
 RN [3]
 RP REVISIONS TO 59-289 (PROTEIN WIS/DISEASE PROTEIN ZUC).
 RP MEDLINE=77021516; PubMed=823945;
 RA Wolfenstein-Todel C., Frangione B., Prelli F., Franklin E.C.;
 RT "The amino acid sequence of 'heavy chain disease' protein ZUC.
 RT Structure of the Fc fragment of immunoglobulin G3.";
 RL Biochem. Biophys. Res. Commun. 71:907-914(1976).
 RN [4]
 RP SEQUENCE FROM N.A. (DISEASE PROTEIN OMM).
 RP MEDLINE=82247835; PubMed=6808505;
 RA Alexander A., Steinmetz M., Barritault D., Frangione B.,
 RA Franklin E.C., Hood L., Buxbaum J.N.;
 RT "Gamma Heavy chain disease in man: cDNA sequence supports partial
 RT gene deletion model.";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:3260-3264(1982).
 CC -!- SUBUNIT: DIMER LINKED BY 12 DISULFIDE BONDS; IT HAS AN EXTRA
 CC INTERCHAIN DISULFIDE BOND AT POSITION 7 IN ADDITION TO THE 11
 CC NORMALLY PRESENT IN THE HINGE REGION
 CC -!- MISCELLANEOUS: THE HEAVY CHAIN DISEASE PROTEIN WIS IS SHOWN.
 CC -!- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 42-76 WAS TAKEN FROM THE
 CC REF. 2.
 CC -!- MISCELLANEOUS: DISEASE PROTEIN WIS IS LACKING MOST OF THE V REGION
 CC AND ALL OF THE CH1 REGION.
 CC -!- MISCELLANEOUS: DISEASE PROTEIN ZUC LACK MOST OF THE V REGION, ALL
 CC OF THE CH1 REGION, AND PART OF THE HINGE COMPARED WITH NORMAL
 CC GAMMA-3 HEAVY CHAINS.
 CC -!- MISCELLANEOUS: DISEASE PROTEIN OMM MAY REPRESENT AN ALLELIC FORM
 CC OR ANOTHER GAMMA CHAIN SUBCLASS.
 CC -!- MISCELLANEOUS: THE HINGE REGION IN GAMMA-3 CHAINS IS ABOUT FOUR
 CC TIMES AS LONG AS IN OTHER GAMMA CHAINS AND CONTAINS THREE
 CC IDENTICAL 15-RESIDUE SEGMENTS PRECEDED BY A SIMILAR 17-RESIDUE
 CC SEGMENT (12-28).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: J00231; AAA52805.1; ALT_SEQ.
 DR PIR: A02149; G3HUWI.
 DR HSSP: P01857; IFCL.
 DR Genew: HGNC:5527; IGHG3.
 DR MIM: 147120;
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003597; Ig_c1.
 DR InterPro: IPR003600; Ig_like.
 DR Pfam: PF00047; Ig; 2.
 DR SMART: SM00410; Ig_like; 1.
 DR SMART: SM00407; Ig_c1; 1.
 DR PROSITE: PS00290; IG_MHC; 1.
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein; Repeat.
 GN DOMAIN 12 73 HINGE.

RT Intrachain disulfide bonds.;
 RL Biochemistry 9:3188-3196(1970).
 [7]
 RP MEDLINE=77070267; PubMed=1002129;
 RX DISULFIDE BONDS.
 RA "Rule of antibody structure. The primary structure of a monoclonal
 RT IgG1 immunoglobulin (myeloma protein Nie), I: Purification and
 RT characterization of the protein, the L- and H-chains, the
 RT cyanogen bromide cleavage products, and the disulfide bridges.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).
 [8]
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
 RA MEDLINE=81208100; PubMed=7236608;
 RX Deisenhofer J.;
 RT "Crystallographic refinement and atomic models of a human Fc fragment
 RT and its complex with fragment B of protein A from Staphylococcus
 RL aureus at 2.9- and 2.8-A resolution.";
 RT Biochemistry 20:2361-2370(1981).
 CC -!- MISCELLANEOUS: NIE HAS THE GLM(17) ALLOTYPIC MARKER, 97-K, & THE
 CC GLM(1) MARKERS, 239-D & 241-L. KOL & EU SEQUENCES HAVE THE GLM(3)
 CC MARKER & THE GLM (NON-1) MARKERS.
 CC -!- MISCELLANEOUS: NIE ALSO DIFFERS IN THE AMIDATION STATES OF
 CC 35,116,198,289 & 272.
 CC -!- MISCELLANEOUS: EU ALSO DIFFERS IN THE AMIDATION STATES OF RESIDUES
 CC 155, 166, 177, 195, 198, 269, AND 272 AND IN THE ORDER OF RESIDUES
 CC 268-272.
 CC -!- MISCELLANEOUS: KOL ALSO DIFFERS IN THE AMIDATION STATES OF
 CC RESIDUES 198,267&272.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J00228; AAC82527.1; ALT_INIT.
 DR PIR; A02146; GHU.
 DR PDB; 1FC1; 15-JUL-92.
 DR PDB; 1FC2; 15-JUL-92.
 DR Genew; HGNC:5525; IGHG1.
 DR MIM; 147100; -
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003597; Ig_c1.
 DR InterPro; IPR003600; Ig_like.
 DR Pfam; PF00047; Ig_3.
 DR SMART; SM00410; Ig_like; 1.
 DR SMART; SM00407; IGHG1; 2.
 DR PROSITE; PS00290; Ig_MHC; 2.
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
 KW 3D-structure.
 FT NON_TER 1 98
 FT DOMAIN 1 98
 FT CH1.
 FT HINGE.
 FT 99 110
 FT DOMAIN 111 223
 FT CH2.
 FT 224 330
 FT DOMAIN 224 330
 FT DISULFID 27 83
 FT DISULFID 103 103
 FT DISULFID 109 109
 FT DISULFID 112 112
 FT DISULFID 144 204
 FT DISULFID 250 308
 FT CARBOHYD 180 180
 FT MOD_RES 330 330
 FT VARIANT 97 97
 FT 239 239
 FT VARIANT 241 241
 FT VARIANT 241 241
 FT STRAND 123 126

FT HELIX 130 134
 FT TURN 136 137
 FT STRAND 141 148
 FT STRAND 158 162
 FT TURN 163 164
 FT STRAND 165 166
 FT STRAND 175 178
 FT STRAND 183 190
 FT HELIX 193 197
 FT TURN 198 199
 FT STRAND 202 206
 FT STRAND 215 219
 FT STRAND 227 227
 FT STRAND 230 234
 FT HELIX 241 242
 FT TURN 241 242
 FT STRAND 245 256
 FT STRAND 260 266
 FT TURN 267 268
 FT STRAND 269 270
 FT STRAND 274 276
 FT STRAND 280 281
 FT TURN 283 284
 FT STRAND 287 296
 FT HELIX 297 301
 FT TURN 302 303
 FT STRAND 306 312
 FT TURN 313 314
 FT TURN 316 317
 FT STRAND 320 324
 SQ SEQUENCE 330 AA; 36106 MW; 3770EE106C2FA33D CRC64;
 Query Match 92.4%; Score 1233; DB 1; Length 330;
 Best Local Similarity 100.0%; Pred. No. 1.1e-34;
 Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 DKHTCPPCPAPELLGGPSVFLFPPPKDRLMISRTPEVTCVVVDVSHEDPEVKFNWYD 61
 Db 104 DKHTCPPCPAPELLGGPSVFLFPPPKDRLMISRTPEVTCVVVDVSHEDPEVKFNWYD 163
 QY 62 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
 Db 164 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 223
 QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 181
 Db 224 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 283
 QY 182 DGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNHYTQKSLSLSPGK 228
 Db 284 DGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNHYTQKSLSLSPGK 330
 RESULT 2
 GC2_HUMAN
 ID GC2_HUMAN STANDARD; PRT; 326 AA.
 AC P01859;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ig gamma-2 chain C region.
 OS IGHG2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE OF 2-326 FROM N.A.
 RX MEDLINE=82197621; PubMed=6804948;
 RA Ellison J.W., Hood L.E.;
 RT "Linkage and sequence homology of two human immunoglobulin gamma
 RT heavy chain constant region genes";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:1984-1988(1982).

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:21 ; Search time 9.40588 seconds
(without alignments)
1084.766 Million cell updates/sec

Title: 2LINK7

Perfect score: 1334

Sequence: 1 MDKTHTCPPCAPPELLGGPS.....GKDLKAFYDKVAEKLEAF 246

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1233	92.4	330	1 GCL_HUMAN	P01857 homo sapien
2	1146	85.9	326	1 GC2_HUMAN	P01859 homo sapien
3	1142	85.6	327	1 GC4_HUMAN	P01861 homo sapien
4	1138	85.3	290	1 GC3_HUMAN	P01860 homo sapien
5	921	69.0	323	1 GC_RABIT	P01870 oryctolagus
6	894	67.0	329	1 GC2_CAVPO	P01862 cavia porce
7	848	63.6	329	1 GC3_MOUSE	P22436 mus musculus
8	840	63.0	333	1 GCB_RAT	P20761 rattus norv
9	838	62.8	398	1 GC3M_MOUSE	P03987 mus musculus
10	825	61.8	326	1 GCL_RAT	P20759 rattus norv
11	823	61.7	324	1 GCL_MOUSE	P01868 mus musculus
12	823	61.7	393	1 GC1M_MOUSE	P01869 mus musculus
13	819	61.4	399	1 GCAM_MOUSE	P01865 mus musculus
14	818	61.3	329	1 GCC_RAT	P20762 rattus norv
15	813	60.9	330	1 GCRA_MOUSE	P01863 mus musculus
16	809	60.6	335	1 GCAB_MOUSE	P01864 mus musculus
17	793	59.4	322	1 GCB_MOUSE	P20760 rattus norv
18	791	59.3	405	1 GCBM_MOUSE	P01867 mus musculus
19	785	58.8	336	1 GCB_MOUSE	P01866 mus musculus
20	387	28.0	421	1 EPC_MOUSE	P06336 mus musculus
21	382	28.6	454	1 MUC_HUMAN	P01871 homo sapien
22	372	27.9	429	1 EPC_RAT	P01855 rattus norv
23	370	27.7	391	1 MUCB_HUMAN	P04220 homo sapien
24	370	27.7	455	1 MUC_MOUSE	P01872 mus musculus
25	369	27.7	428	1 EPC_HUMAN	P01854 homo sapien
26	367	27.5	458	1 MUC_RABIT	P03988 oryctolagus
27	364	27.3	476	1 MUCM_MOUSE	P01873 mus musculus
28	361	27.1	479	1 MUCM_RABIT	P04221 oryctolagus
29	356	26.7	454	1 MUC_MESAU	P06337 mesocricetu
30	349	26.2	450	1 MUC_CANFA	P01874 canis fami
31	347	26.0	457	1 MUC_SUNMU	P20768 suncus muri
32	326	24.4	446	1 MUC_CHICK	P01875 gallus gall
33	321	24.1	438	1 HVC2_HETFR	P23085 heterodontu

ALIGNMENTS

RESULT 1

ID	GCL_HUMAN	STANDARD;	PRT;	330 AA.
AC	P01857:			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Ig gamma-1 chain C region.			
GN	IGHG1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=82274238; PubMed=6287432;			
RA	Ellison J.W., Berson B.J., Hood L.E.;			
RT	"The nucleotide sequence of a human Immunoglobulin C gamma gene."			
RL	Nucleic Acids Res. 10:4071-4079(1982).			
RN	[2]			
RP	SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).			
RX	MEDLINE=71064024; PubMed=5489771;			
RA	Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,			
RA	Waxdal M.J., Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. VII. Amino			
RL	acid sequence of heavy-chain cyanogen bromide fragments H1-H4."			
RN	Biochemistry 9:3161-3170(1970).			
RP	SEQUENCE OF 136-329 (EU).			
RX	MEDLINE=71064025; PubMed=5530842;			
RA	Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,			
RA	Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. 8. Amino			
RL	acid sequence of heavy-chain cyanogen bromide fragments H5-H7."			
RN	Biochemistry 9:3171-3181(1970).			
RP	SEQUENCE (MYELOMA PROTEIN NIE).			
RX	MEDLINE=77070269; PubMed=826475;			
RA	Ponstingl H., Hilschmann N.;			
RT	"The rule of antibody structure. The primary structure of a			
RL	monoclonal IgG1 immunoglobulin (myeloma protein NIE). III. The			
RT	chymotryptic peptides of the H-chain, alignment of the tryptic			
RL	peptides and discussion of the complete structure."			
RN	Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).			
RP	SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.			
RX	MEDLINE=83289131; PubMed=6884994;			
RA	Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;			
RT	"Three-dimensional structure determination of antibodies. Primary			
RL	structure of crystallized monoclonal immunoglobulin IgG1 KOL, I."			
RN	Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).			
RP	DISULFIDE BONDS.			
RX	MEDLINE=71064027; PubMed=4923144;			
RA	Gall W.E., Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. X.			

34	317	23.8	6486	1 TYCC_BACBR	O30409 b tyrocidin
35	313	23.5	438	1 HVCS_HETFR	P23087 heterodontu
36	312	23.4	5037	1 RYRL_RABIT	P11716 oryctolagus
37	312	23.4	5038	1 RYRL_HUMAN	P21817 homo sapien
38	311	23.3	461	1 HVCM_HETFR	P23088 heterodontu
39	310	23.2	4969	1 RYR2_RABIT	P30957 oryctolagus
40	308	23.1	299	1 ALC_RABIT	P01879 oryctolagus
41	307	23.0	4725	1 DYHC_DICDI	P34036 dictyosteli
42	306	22.9	4367	1 DYHC_NEUCR	P45443 neurospora
43	306	22.9	4967	1 RYR2_HUMAN	Q92736 homo sapien
44	305	22.9	370	1 HVC1_HETFR	P23084 heterodontu
45	305	22.9	4393	1 PGBM_HUMAN	P98160 homo sapien

RESULT 14

G2GP

Ig gamma-2 chain C region - guinea pig

C:Species: Cavia porcellus (guinea pig)

C:Date: 07-May-1981 #sequence-revision 07-May-1981 #text-change 16-Jul-1999

C:Accession: A94553; A90352; A90359; A90384; A90385; A02151

R:Trischmann, T.M.

submitted to the Atlas, April 1975

A:Reference number: A94553

A:Accession: A94553

A:Molecule type: protein

A:Residues: 1-3 <TRI>

R:Rirshtein, B.K.; Hussain, Q.Z.; Cebra, J.J.

Biochemistry 10, 18-25, 1971

A:Title: Structure of heavy chain from strain 13 guinea pig immunoglobulin-G(2). III. An

A:Reference number: A90352; MUID:71058471; PMID:5538606

A:Accession: A90352

A:Molecule type: protein

A:Residues: 4-68 <BIR>

R:Turner, K.J.; Cebra, J.J.

Biochemistry 10, 9-17, 1971

A:Title: Structure of heavy chain from strain 13 guinea pig immunoglobulin-G(2). II. An

A:Reference number: A90359; MUID:71058486; PMID:5538616

A:Accession: A90359

A:Molecule type: protein

A:Residues: 69-133;312-329 <TUR>

R:Tracey, D.E.; Cebra, J.J.

Biochemistry 13, 4796-4803, 1974

A:Title: Primary structure of the C-H2 homology region from guinea pig IgG2 antibodies.

A:Reference number: A90384; MUID:75036072; PMID:4429665

A:Accession: A90384

A:Molecule type: protein

A:Residues: 134-226 <TRA>

R:Trischmann, T.M.; Cebra, J.J.

Biochemistry 13, 4804-4811, 1974

A:Title: Primary structure of the C-H3 homology region from guinea pig IgG2 antibodies.

A:Reference number: A90385; MUID:75036073; PMID:4609467

A:Accession: A90385

A:Molecule type: protein

A:Residues: 227-311 <TR2>

R:Oliveira, B.; Lamm, M.E.

Biochemistry 10, 26-31, 1971

A:Title: Interchain disulfide bridges of guinea pig gamma-2 - immunoglobulin.

A:Reference number: A90354; MUID:71058474; PMID:4922544

A:Contents: annotation; disulfide bonds

A:Note: Cys-16 is involved in a heavy-light chain bond

A:Note: Cys-105, Cys-107, and Cys-110 form inter-heavy chain bonds

C:Comment: This chain was isolated from pooled serum of strain 13 inbred guinea pigs.

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap

chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin

F:21-81/Domain: immunoglobulin homology <IM1>

F:135-204/Domain: immunoglobulin homology <IM2>

F:241-310/Domain: immunoglobulin homology <IM3>

F:28-79/Disulfide bonds: #status experimental

F:142-202/Disulfide bonds: #status experimental

F:178/Binding site: carbonylate (Asn) (covalent) #status experimental

F:248-308/Disulfide bonds: #status experimental

Query Match 67.0%; Score 894; DB 1; Length 329;

Best Local Similarity 68.8%; Pred. No. 8.9e-23;

Matches 165; Conservative 25; Mismatches 37; Indels 13; Gaps 2;

QY 1 MKT-----HTCPCPAPELLGSPVFLFPKPKDTLMISRTPEVTCVVVDVSH 49

:||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

DB 90 VDKTVEFIRTPBPBCTCPKPPENLGGPSVFIFPPKPKDTLMISLTPEVTCVVVDVSQ 149

QY 50 EDPEVKFNYYVDGVEVHNNAKTPREFQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKAL 109

:|||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||

DB 150 DEPEVQTFWFDVNKNPGVNAETPRVQYNTTFRVESVLPFIQHODWLRGKEFKCKVYNKAL 209

QY 110 PAPIEKTISKAKGQPREPOVYITPPSRDELITNQVSLTCLVKGFYPSDIAVEWESNGQP- 168

Db 151 FTWYINNEQVARTPLRQQQFNSTIRVSTLTPIHQDWLRGKFKCKVHNKALPAPIEK 210
QY 116 TISKAGQPREQVYTLPLPSRDELTKNOVSLTCLVKGFVPSDIAVESNGOPENNYKTT 175
Db 211 TISKAGQPLEPKVYTMGPPEELSSRSVSLTCMNGFFPSDISVEWENKRAEDNYKTT 270
QY 176 PVLSDSGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSLSPGK 228
Db 271 PAVLSDSGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSRSPGK 323

RESULT 11

I47160
Ig gamma 2b chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47160
R:Kaskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a
A:Reference number: I47158; MUID:95015845; PMID:7930579
A:Accession: I47160
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-328 <KAC>
A:Cross-references: EMBL:U03780; NID:9433125; PIDN:AAA52218.1; PID:9433126
C:Genetics:
A:Gene: IgG2b
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 68.7%; Score 917; DB 2; Length 328;
Best Local Similarity 65.4%; Pred. No. 1.4e-23;
Matches 166; Conservative 32; Mismatches 24; Indels 32; Gaps 4;

QY 3 KTHTC-----PPCP-----APELLGGPSVFLPPPKPKDTLMISR 36
Db 79 KSYTCNVNHPATTTKVDRKVGTKTPPCPCIPACESP-----GPSVFIFPPPKPKDTLMISR 134
QY 37 TPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNKATKPREQYNSTYRVVSVLTVLHQDWLNL 96
Db 135 TPQVTCVVVDVSHEDPEVKFNMYVDGVEVHNKATKPREQYNSTYRVVSVLTVLHQDWLNL 194
QY 97 GREYCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPS 156
Db 195 GREYCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPS 254
QY 157 DIAVESNGQ--PENNYKTTTPVLSDSGSFFLYSKLTVDKSRWQGNVFCSCVMHEALH 214
Db 255 DIDVEWQRNGQPEPEGNRYTTPQDDVDGTYFLYSKFSVDKASWQGGGIFQCAVMHEALH 314
QY 215 NHYTKSLSLSPGK 228
Db 315 NHYTKSLSLSPGK 328

RESULT 12

I47159
Ig gamma 2a chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47159
R:Kaskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of a
A:Reference number: I47158; MUID:95015845; PMID:7930579
A:Accession: I47159
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-328 <KAC>
A:Cross-references: EMBL:U03779; NID:9433123; PIDN:AAA52217.1; PID:9433124
C:Genetics:
A:Gene: IgG2a

C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:133-202/Domain: immunoglobulin homology <IMM>

Query Match 68.7%; Score 917; DB 2; Length 328;
Best Local Similarity 65.4%; Pred. No. 1.4e-23;
Matches 166; Conservative 32; Mismatches 24; Indels 32; Gaps 4;

QY 3 KTHTC-----PPCP-----APELLGGPSVFLPPPKPKDTLMISR 36
Db 79 KSYTCNVNHPATTTKVDRKVGTKTPPCPCIPACESP-----GPSVFIFPPPKPKDTLMISR 134
QY 37 TPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNKATKPREQYNSTYRVVSVLTVLHQDWLNL 96
Db 135 TPQVTCVVVDVSHEDPEVKFNMYVDGVEVHNKATKPREQYNSTYRVVSVLTVLHQDWLNL 194
QY 97 GREYCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPS 156
Db 195 GREYCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPS 254
QY 157 DIAVESNGQ--PENNYKTTTPVLSDSGSFFLYSKLTVDKSRWQGNVFCSCVMHEALH 214
Db 255 DIDVEWQRNGQPEPEGNRYTTPQDDVDGTYFLYSKFSVDKASWQGGGIFQCAVMHEALH 314
QY 215 NHYTKSLSLSPGK 228
Db 315 NHYTKSLSLSPGK 328

RESULT 13

I47162
Ig gamma 4 chain constant region - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 21-Jan-2000
C:Accession: I47162
R:Kaskovics, I.; Sun, J.; Butler, J.E.
J. Immunol. 153, 3565-3573, 1994
A:Title: Five putative subclasses of swine IgG identified from the cDNA sequences of
A:Reference number: I47158; MUID:95015845; PMID:7930579
A:Accession: I47162
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-277 <KAC>
A:Cross-references: EMBL:U03782; NID:9433129; PIDN:AAA52220.1; PID:9433130
C:Genetics:
A:Gene: IgG4
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:82-151/Domain: immunoglobulin homology <IMM>

Query Match 68.3%; Score 911; DB 2; Length 277;
Best Local Similarity 66.4%; Pred. No. 1.3e-23;
Matches 166; Conservative 31; Mismatches 29; Indels 24; Gaps 4;

QY 3 KTHTC-----PPCP-----APELLG-GPSVFLPPPKPKDTLMISRTPEV 40
Db 28 KSYTCNVNHPATTTKVDRKVGTKTPPCPCIPACESP-----GPSVFIFPPPKPKDTLMISRTPKV 87
QY 41 TCVVVDVSHEDPEVKFNMYVDGVEVHNKATKPREQYNSTYRVVSVLTVLHQDWLNGKEY 100
Db 88 TCVVVDVSHEDPEVKFNMYVDGVEVHNKATKPREQYNSTYRVVSVLTVLHQDWLNGKEY 147
QY 101 KCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAV 160
Db 148 KCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAV 207
QY 161 EWESNGQ--PENNYKTTTPVLSDSGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYT 218
Db 208 EWQRNGQPEPEGNRYTTPQDDVDGTYFLYSKFSVDKASWQGGGIFQCAVMHEALHNHYT 267
QY 219 QKSLSLSPGK 228
Db 268 QKSLSLSPGK 277

A:Residues: 1-330 <ELL>
A:Cross-references: EMBL:Z17370
A:Note: This sequence has the G1m(17) allotypic marker, 97-Lys, and the G1m(1) markers,
R: Haggis, L.J.
submitted to the EMBL Data Library, October 1992
A:Reference number: S33904
A:Accession: S36861
A:Molecule type: DNA
A:Residues: 2-330 <HAR>
A:Cross-references: EMBL:Z17370
R:Takahashi, N.; Ueda, S.; Obata, M.; Nikaido, T.; Nakai, S.; Honjo, T.
Cell, 29, 671-679, 1982
A:Title: Structure of human immunoglobulin gamma genes: Implications for evolution of a
A:Reference number: S33887; MUID:83001943; PMID:6811139
A:Accession: S33887
A:Molecule type: DNA
A:Residues: 88-113; 235-330 <TAK>
A:Cross-references: EMBL:Z17370
R:Cunningham, B.A.; Rutishauser, U.; Gall, W.E.; Gottlieb, P.D.; Waxdal, M.J.; Edelman,
Biochemistry 9, 3161-3170, 1970
A:Title: The covalent structure of a human gammaG-immunoglobulin. VII. Amino acid sequen
A:Reference number: A90563; MUID:71064024; PMID:5489771
A:Contents: myeloma protein Eu
A:Accession: B90563
A:Molecule type: protein
A:Residues: 1-96, 'R', 98-135 <CUN>
A:Note: This sequence has the G1m(3) marker, 97-Arg
R: Rutishauser, U.; Cunningham, B.A.; Bennett, C.; Konigsberg, W.H.; Edelman, G.M.
Biochemistry 9, 3171-3181, 1970
A:Title: The covalent structure of a human gammaG-immunoglobulin. VIII. Amino acid sequen
A:Reference number: A90564; MUID:71064025; PMID:5530842
A:Contents: Eu
A:Accession: A90564
A:Molecule type: protein
A:Residues: 136-154, 'Q', 156-165, 'Q', 167-176, 'Q', 178-194, 'N', 196-197, 'D', 199-238, 'E', 240,
A:Note: This sequence has the G1m(non-1) markers, 239-Glu and 241-Met
R: Ponting, H.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 357, 1571-1604, 1976
A:Title: Die Primaerstruktur eines monoklonalen IgG1-Immunglobulins (Myelomprotein Nile)
igen Primaerstruktur.
A:Reference number: A91668; MUID:77070269; PMID:826475
A:Contents: myeloma protein Nie
A:Accession: B91668
A:Molecule type: protein
A:Residues: 1-34, 'Q', 36-96, 'K', 98-115, 'Q', 117-197, 'D', 199-238, 'D', 240, 'L', 242-268, 'E', 27
A:Note: This sequence has the G1m(17) and G1m(1) markers
R: Schmidt, W.E.; Jung, H.D.; Palm, W.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 364, 713-747, 1983
A:Title: Die Primaerstruktur des kristallisierten monoklonalen Immunglobulins IgG1 KOL
A:Reference number: A91723; MUID:83289131; PMID:6884994
A:Contents: myeloma protein KOL; disulfide bonds
A:Accession: A91723
A:Molecule type: protein
A:Residues: 1-96, 'R', 98-197, 'D', 199-238, 'E', 240, 'M', 242-266, 'D', 268-271, 'D', 273-330 <SCH
A:Note: This sequence has the G1m(3) and G1m(non-1) markers
R: Gall, W.E.; Edelman, G.M.
Biochemistry 9, 3188-3196, 1970
A:Title: The covalent structure of a human gammaG-immunoglobulin. X. Intrachain disulfid
A:Reference number: A90565; MUID:71064027; PMID:4923144
A:Contents: annotation; disulfide bonds
R: Draker, L.; Schwarz, J.; Reichel, W.; Hilschmann, N.
Hoppe-Seyler's Z. Physiol. Chem. 357, 1515-1540, 1976
A:Title: Rule of antibody structure. The primary structure of monoclonal IgG1 immunoglob
enbromide cleavage products, and the disulfide bridges.
A:Reference number: A91667; MUID:77070267; PMID:1002129
A:Contents: annotation; disulfide bonds
A:Accession: B91667
A:Gene: GDB:IGHG1
A:Cross-references: GDB:120085; OMIM:147100
A:Map position: 14q32.33-14q32.33
A:Introns: 99/1; 114/1; 224/1
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap

chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:20-85/Domain: immunoglobulin homology <IM1>
F:137-206/Domain: immunoglobulin homology <IM2>
F:243-310/Domain: immunoglobulin homology <IM3>
F:27-83,144-204,250-308/disulfide bonds: #status experimental
F:103/disulfide bonds: interchain (to light chain) #status experimental
F:109,112/disulfide bonds: interchain (to heavy chain) #status experimental
F:180/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 92.48; Score 1233; DB 1; Length 330;
Best Local Similarity 100.0%; Pred. No. 2.3e-34;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DKTHCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
|||||
Db 104 DKTHCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 163
|||||

Qy 62 GVEVHNATKPREQYNSTYRVVSVLTVLDHQLNGKEYCKVSNKALPAPIEKTISKAK 121
|||||
Db 164 GVEVHNATKPREQYNSTYRVVSVLTVLDHQLNGKEYCKVSNKALPAPIEKTISKAK 223
|||||

Qy 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 181
|||||
Db 224 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 283
|||||

Qy 182 DGSFFLYSKLVDRKRWQGVNFCVSMHEALHNHYTOKSLSLSPGK 228
|||||
Db 284 DGSFFLYSKLVDRKRWQGVNFCVSMHEALHNHYTOKSLSLSPGK 330
|||||

RESULT 3
S69339
Ig heavy chain V region precursor - human
C:Species: Homo sapiens (man)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Dec-2000
C:Accession: S69339; S72664
R:Khamilichi, A.A.; Aucouturier, P.; Preud'homme, J.L.; Cogne, M.
Eur. J. Biochem. 229, 54-60, 1995
A:Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease.
A:Reference number: S69339; MUID:95262687; PMID:7744049
A:Accession: S69339
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-374 <KHA>
A:Cross-references: EMBL:X81695
R:Khamilichi, A.A.
submitted to the EMBL Data Library, September 1994
A:Reference number: S72664
A:Accession: S72664
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-140, 'C', 142-374 <KH2>
A:Cross-references: EMBL:X81695
C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 92.08; Score 1227; DB 2; Length 374;
Best Local Similarity 99.18; Pred. No. 5.5e-34;
Matches 225; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DKTHCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
|||||
Db 148 DKTHCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 207
|||||

Qy 62 GVEVHNATKPREQYNSTYRVVSVLTVLDHQLNGKEYCKVSNKALPAPIEKTISKAK 121
|||||
Db 208 GVEVHNATKPREQYNSTYRVVSVLTVLDHQLNGKEYCKVSNKALPAPIEKTISKAK 267
|||||

Qy 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 181
|||||
Db 268 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 327
|||||

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:26 ; Search time 18.0882 Seconds
(without alignments)
1307.428 Million cell updates/sec

Title: 2LINK7
Perfect score: 1334
Sequence: 1 MDKTHTCPPCPAPELGGPS.....GKDWLKAFYDKVAEKLKEAF 246

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.0

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1233	92.4	255	4	Ig gamma-1 chain C
2	1233	92.4	330	1	Ig gamma-1 chain C
3	1227	92.0	374	2	Ig heavy chain V r
4	1180	88.5	234	2	Ig gamma chain C r
5	1158	86.8	377	2	Ig gamma-3 chain C
6	1156	86.7	377	2	Ig gamma-3 chain C
7	1146	85.9	326	1	Ig gamma-2 chain C
8	1142	85.6	327	1	Ig gamma-4 chain C
9	1133	84.9	289	1	Ig gamma-3 heavy c
10	921	69.0	323	1	Ig gamma chain C r
11	917	68.7	328	2	Ig gamma 2b chain
12	917	68.7	328	2	Ig gamma 2a chain
13	911	68.3	277	2	Ig gamma-4 chain c
14	894	67.0	329	1	Ig gamma-2 chain C
15	888	66.6	328	2	Ig gamma 1 chain C
16	885	66.3	328	2	Ig gamma 3 chain c
17	860	64.5	470	2	Ig heavy chain pre
18	858	64.3	472	2	Ig gamma-1 chain -
19	853	63.9	308	2	Ig heavy chain C r
20	848	63.6	329	1	Ig gamma-3 chain C
21	840	63.0	333	2	Ig gamma-2b chain
22	838	62.8	398	1	Ig gamma-3 chain C
23	834	62.5	444	2	monoclonal antibod
24	825	61.8	326	2	Ig gamma-1 chain C
25	823	61.7	324	1	Ig gamma-1 chain C
26	823	61.7	393	1	Ig gamma-1 chain C
27	819	61.4	399	1	Ig gamma-2a chain
28	813	61.3	329	2	Ig gamma-2c chain
29	813	60.9	330	1	Ig gamma-2a chain

30	813	60.9	469	2	S37483	Ig gamma-2a chain
31	809	60.6	335	1	G2MSAB	Ig gamma-2a chain
32	799	59.9	446	2	S40295	Ig gamma-2a chain
33	793	59.4	322	2	PS0019	Ig gamma-2a chain
34	791	59.3	405	1	G2MSBM	Ig gamma-2b chain
35	787	59.0	474	1	G2MS11	Ig gamma-2b chain
36	774	58.0	327	2	S06611	Ig gamma-2 chain C
37	766	57.4	475	2	S01321	Ig gamma-2b chain
38	707	53.0	180	2	I46732	Ig gamma heavy cha
39	602	45.1	218	2	A36040	Ig heavy chain V-I
40	601	45.1	249	2	S69340	Ig heavy chain V-I
41	571	42.8	152	2	S14236	Ig gamma-1 chain C
42	466	34.9	26926	1	I38344	titin, cardiac mus
43	437	32.8	572	2	B46529	Ig Y heavy chain (
44	422	31.6	548	2	S38864	Ig epsilon chain C
45	407	30.5	448	2	S03186	Ig heavy chain C r

ALIGNMENTS

RESULT 1

S31866

Ig gamma-1 chain C region - synthetic

C:Species: synthetic

A:Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli

C:Date: 06-Jan-1995 #sequence_revision 17-Mar-1997 #text_change 19-May-2000

C:Accession: S31866

R:Filipula, D.

submitted to the EMBL Data Library, February 1993

A:Description: Screening method for protein-protein interactions of cloned gene products

A:Reference number: S31866

A:Accession: S31866

A:Molecule type: mRNA

A:Residues: 1-255 <FILL>

A:Cross-references: EMBL:X70421; NID:g33068; PIDN:CAA49866.1; PID:g33069

C:Keywords: immunoglobulin

F:1-22/Region: Escherichia coli outer membrane protein A precursor

F:23-255/Region: human Ig gamma-1 chain C region

Query Match

Best Local Similarity 92.4%; Score 1233; DB 4; Length 255;

Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	2	DKTHTCPPCPAPELGGPSVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVD	61
DB	29	DKTHTCPPCPAPELGGPSVFLFPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVD	88
QY	62	GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK	121
DB	89	GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK	148
QY	122	GGPREPQVYTLPPSRDELTKNQVSLTCLVKGPYPSDIAVEWESNGQPENNYKTPPVLD	181
DB	149	GGPREPQVYTLPPSRDELTKNQVSLTCLVKGPYPSDIAVEWESNGQPENNYKTPPVLD	208
QY	182	DGSFFLYSKLTVDKSRWQGNVFSCVMEALHNHYTKSLSPGK	228
DB	209	DGSFFLYSKLTVDKSRWQGNVFSCVMEALHNHYTKSLSPGK	255

RESULT 2

GHU

Ig gamma-1 chain C region - human

C:Species: Homo sapiens (man)

C:Date: 31-Jan-1981 #sequence_revision 18-Aug-1982 #text_change 16-Jul-1999

C:Accession: A93433; S36861; S33887; B90563; A90564; B91668; A91723; A02146

R:Ellison, J.W.; Berson, B.J.; Hood, L.E.

Nucleic Acids Res. 10, 4071-4079, 1982

A:Title: The nucleotide sequence of a human immunoglobulin C-gamma gene.

A:Reference number: A93433; MUID:82274238; PMID:6287432

A:Accession: A93433

A:Molecule type: DNA

Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	
Qy 1	MDKTHTCPPCPAPPELLGGPSVLEFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60
Db :	
Db 2	VDKTHTCPPCPAPPELLGGPSVLEFPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 61
Qy 61	DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
Db :	
Db 62	DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 121
Qy 121	KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGYFSPDIAVEWESNGQPENNYKTTTPPVL 180
Db :	
Db 122	KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGYFSPDIAVEWESNGQPENNYKTTTPPVL 181
Qy 181	SDGSFFLYSKLTVDKSRWQOGNPFSCVMHEALHNHYTQKSLSLSPGK 228
Db :	
Db 182	SDGSFFLYSKLTVDKSRWQOGNPFSCVMHEALHNHYTQKSLSLSPGK 229

Search completed: April 21, 2003, 10:51:15
Job time : 13.3353 secs

```
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFg1 heavy chain - DNase I fusion
US-09-825-012-49

Query Match          92.7%; Score 1237; DB 10; Length 730;
Best Local Similarity 48.7%; Pred. No. 6.4e-32;
Matches 233; Conservative 6; Mismatches 5; Indels 234; Gaps 4;

QY  2 DKHTCCPPCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
Db  241 DKHTCCPPCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 300

QY  62 GVEVHNAKTPREEQNSYRYVSVLTVLHQDWLNKKEYCKVSNKALPAPIETISKAK 121
Db  301 GVEVHNAKTPREEQNSYRYVSVLTVLHQDWLNKKEYCKVSNKALPAPIETISKAK 360

QY  122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 181
Db  361 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 420

QY  182 DGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNNHYTKQSLSLSP----- 226
Db  421 DGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNNHYTKQSLSLSP----- 480

QY  227 ----- 226
Db  481 FGETKSNATLVSIIYOILSRVDIALVQEVDRSHLTAVGKLLDNLQADPDTHYHYVSEP 540

QY  227 -----GKD----- 229
Db  541 LGRNSYKERYLFVYRPDQVSAVDYDDGCEPCGNDTENREPAIVRFFSRFTEVREFAI 600

QY  230 ----- 229
Db  601 VPLHAAPGDAVAEIDALDYLDVQEKWGLEDMVLMGDFNAGCSYVRPSQWSSIRLWTS 660

QY  230 ---WL-----KAFYDKV-----AEKLEA 245
Db  661 TFQWLIPDSADTATTHCAIDRVVAGMLLRGAVVPSALPFPFOAAYGLSDQLAQ 718

RESULT 14
US-09-825-012-58
; Sequence 58, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 58
; LENGTH: 740
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFg1 heavy chain - DNase I fusion
US-09-825-012-58

Query Match          92.7%; Score 1237; DB 10; Length 740;
Best Local Similarity 48.7%; Pred. No. 6.6e-32;
Matches 233; Conservative 6; Mismatches 5; Indels 234; Gaps 4;

QY  2 DKHTCCPPCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 61
Db  241 DKHTCCPPCAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 300

QY  62 GVEVHNAKTPREEQNSYRYVSVLTVLHQDWLNKKEYCKVSNKALPAPIETISKAK 121
Db  301 GVEVHNAKTPREEQNSYRYVSVLTVLHQDWLNKKEYCKVSNKALPAPIETISKAK 360

QY  122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 181
Db  361 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 420

QY  182 DGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNNHYTKQSLSLSP----- 226
Db  421 DGSFFLYSKLTVDKSRWQQGNVSCVMHEALHNNHYTKQSLSLSP----- 480

QY  227 ----- 226
Db  481 FGETKSNATLVSIIYOILSRVDIALVQEVDRSHLTAVGKLLDNLQADPDTHYHYVSEP 540

QY  227 -----GKD----- 229
Db  541 LGRNSYKERYLFVYRPDQVSAVDYDDGCEPCGNDTENREPAIVRFFSRFTEVREFAI 600

QY  230 ----- 229
Db  601 VPLHAAPGDAVAEIDALDYLDVQEKWGLEDMVLMGDFNAGCSYVRPSQWSSIRLWTS 660

QY  230 ---WL-----KAFYDKV-----AEKLEA 245
Db  661 TFQWLIPDSADTATTHCAIDRVVAGMLLRGAVVPSALPFPFOAAYGLSDQLAQ 718

RESULT 15
US-10-215-297-2
; Sequence 2, Application US/10215297
; Publication No. US20020192222A1
; GENERAL INFORMATION:
; APPLICANT: Blumberg, Richard S.
; APPLICANT: Simister, Neil E.
; APPLICANT: Lencer, Wayne I.
; TITLE OF INVENTION: RECEPTOR SPECIFIC TRANSEPITHELIAL TRANSPORT OF
; TITLE OF INVENTION: THERAPEUTICS
; FILE REFERENCE: S1383/77003
; CURRENT APPLICATION NUMBER: US/10/215,297
; CURRENT FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: US 08/578,171
; PRIOR FILING DATE: 1995-12-29
; PRIOR APPLICATION NUMBER: US 08/374,159
; PRIOR FILING DATE: 1995-01-17
; PRIOR APPLICATION NUMBER: US 09/122,144
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 22
; SEQ ID NO 2
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION:
US-10-215-297-2

Query Match          92.5%; Score 1234; DB 9; Length 229;
Best Local Similarity 99.6%; Pred. No. 5.5e-33;
```

;; PRIOR FILING DATE: 2000-05-03
;; NUMBER OF SEQ ID NOS: 135
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 109
;; LENGTH: 282
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; OTHER INFORMATION: Echistatin Fc-peptide
;; NAME/KEY: misc_feature
;; LOCATION: (1)..(1)
;; OTHER INFORMATION: NdeI site
;; NAME/KEY: misc_feature
;; LOCATION: (854)..(854)
;; OTHER INFORMATION: BamHI site
US-09-840-277-109

Query Match 92.8%; Score 1238; DB 9; Length 282;
Best Local Similarity 100.0%; Pred. No. 6.e-33;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKHTCPCPAPPELLGGPSVFLFPPPKDPTLMISRPEVTCVVVDVSHEDPEVKFNWYV 60
Db 1 MDKHTCPCPAPPELLGGPSVFLFPPPKDPTLMISRPEVTCVVVDVSHEDPEVKFNWYV 60
QY 61 DGEVHNAKTPREEQNSTYRVSVLTQLHQLDNLNGKEYCKVSNKALPAPIEKTISK 120
Db 61 DGEVHNAKTPREEQNSTYRVSVLTQLHQLDNLNGKEYCKVSNKALPAPIEKTISK 120
QY 121 KGPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 180
Db 121 KGPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 180
QY 181 SDGSFFLYSKLTVDKSRWQGNVFSCSVMHLEALHNHYTQKSLSLSPGK 228
Db 181 SDGSFFLYSKLTVDKSRWQGNVFSCSVMHLEALHNHYTQKSLSLSPGK 228

RESULT 11
US-09-784-623-14
;; Sequence 14, Application US/09784623
;; Patent No. US20020009454A1
;; GENERAL INFORMATION:
;; APPLICANT: Boone, Thomas C.
;; APPLICANT: Hersenson, Susan
;; APPLICANT: Bevilacqua, Michael P.
;; APPLICANT: Collins, David S.
;; TITLE OF INVENTION: COMPOSITION AND METHOD FOR TREATING INFLAMMATORY
;; FILE REFERENCE: A-365F
;; CURRENT APPLICATION NUMBER: US/09/784,623
;; CURRENT FILING DATE: 2001-02-15
;; PRIOR APPLICATION NUMBER: 09/131,247
;; PRIOR FILING DATE: 1998-08-07
;; PRIOR APPLICATION NUMBER: PCT/US 97/02131
;; PRIOR FILING DATE: 1997-02-10
;; NUMBER OF SEQ ID NOS: 16
;; SOFTWARE: PatentIn ver. 2.0
;; SEQ ID NO 14
;; LENGTH: 389
;; TYPE: PRT
;; ORGANISM: Human
US-09-784-623-14

Query Match 92.8%; Score 1238; DB 10; Length 389;
Best Local Similarity 72.7%; Pred. No. 1.4e-32;
Matches 234; Conservative 4; Mismatches 5; Indels 79; Gaps 4;

QY 2 DKHTCPCPAPPELLGGPSVFLFPPPKDPTLMISRPEVTCVVVDVSHEDPEVKFNWYV 61
Db 10 DKHTCPCPAPPELLGGPSVFLFPPPKDPTLMISRPEVTCVVVDVSHEDPEVKFNWYV 69
QY 62 GVEVHNAKTPREEQNSTYRVSVLTQLHQLDNLNGKEYCKVSNKALPAPIEKTISK 121

Db 70 GVEVHNAKTPREEQNSTYRVSVLTQLHQLDNLNGKEYCKVSNKALPAPIEKTISK 129
QY 122 GOPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 181
Db 130 GOPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 189
QY 182 DGSEFFLYSKLTVDKSRWQGNVFSCSVMHLEALHNHYTQKSLSLSPGK----- 228
Db 190 DGSEFFLYSKLTVDKSRWQGNVFSCSVMHLEALHNHYTQKSLSLSPGKMKRSPGRKSSKMQA 249
QY 229 --DW---LKAFY-----DKV----- 238
Db 250 FRIWDVNAKTFYLRNNOLVAGYLQGNVNLEEKIDVVPPIEPHALFLGIHGKMLSCVKS 309
QY 239 -----AEKLE 244

Db 310 GDETRLOLEAVNITDLSNRKQ 331
RESULT 12
US-09-854-864-31
;; Sequence 31, Application US/09854864
;; Patent No. US20020081296A1
;; GENERAL INFORMATION:
;; APPLICANT: THEILL, LARS EYDE
;; APPLICANT: YU, GANG
;; TITLE OF INVENTION: METHODS AND COMPOSITIONS OF MATTER CONCERNING APRIL/G70, BCMA
;; FILE REFERENCE: BLYS/AGP-3, AND TACI
;; CURRENT APPLICATION NUMBER: US/09/854,864
;; CURRENT FILING DATE: 2001-09-11
;; PRIOR APPLICATION NUMBER: US 60/204,039
;; PRIOR FILING DATE: 2000-05-12
;; PRIOR APPLICATION NUMBER: US 60/214,591
;; PRIOR FILING DATE: 2000-06-27
;; NUMBER OF SEQ ID NOS: 31
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 31
;; LENGTH: 394
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-854-864-31

Query Match 92.7%; Score 1237; DB 10; Length 394;
Best Local Similarity 78.6%; Pred. No. 1.5e-32;
Matches 231; Conservative 1; Mismatches 5; Indels 57; Gaps 2;

QY 2 DKHTCPCPAPPELLGGPSVFLFPPPKDPTLMISRPEVTCVVVDVSHEDPEVKFNWYV 61
Db 20 DKHTCPCPAPPELLGGPSVFLFPPPKDPTLMISRPEVTCVVVDVSHEDPEVKFNWYV 79
QY 62 GVEVHNAKTPREEQNSTYRVSVLTQLHQLDNLNGKEYCKVSNKALPAPIEKTISK 121
Db 80 GVEVHNAKTPREEQNSTYRVSVLTQLHQLDNLNGKEYCKVSNKALPAPIEKTISK 139
QY 122 GOPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 181
Db 140 GOPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 199
QY 182 DGSEFFLYSKLTVDKSRWQGNVFSCSVMHLEALHNHYTQKSLSLSPGK----- 229
Db 200 DGSEFFLYSKLTVDKSRWQGNVFSCSVMHLEALHNHYTQKSLSLSPGKSRAVLTKQKKQH 259
QY 230 -----WLKA-----FYDKV 238
Db 260 SVLHLVPINATSKDDSDVTEVMQPALRRGRGLQAQGYGVRIQDAGVYLLYSQV 313

RESULT 13
US-09-825-012-49
;; Sequence 49, Application US/09825012
;; Patent No. US2002012798A1

```
; Sequence 2, Application US/09840669B
; Publication No. US20030040470A1
; GENERAL INFORMATION:
; APPLICANT: KOHNO, TADAHIKO
; TITLE OF INVENTION: APO-AI/II PEPTIDE DERIVATIVES
; FILE REFERENCE: A-690
; CURRENT APPLICATION NUMBER: US/09/840,669B
; CURRENT FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/198,920
; PRIOR FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-840-669B-2

Query Match          92.8%; Score 1238; DB 9; Length 228;
Best Local Similarity 100.0%; Pred. No. 4e-33;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
DB 1 MDKTHCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
QY 61 DGVEVHNAKTKPREEQNTSYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKA 120
DB 61 DGVEVHNAKTKPREEQNTSYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKA 120
QY 121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 180
DB 121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 180
QY 181 SDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNYHTOKSLSLSPGK 228
DB 181 SDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNYHTOKSLSLSPGK 228

RESULT 9
US-09-847-712-2
; Sequence 2, Application US/09847712
; Patent No. US20020090646A1
; GENERAL INFORMATION:
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: MARSHALL, WILLIAM S.
; APPLICANT: REYNOLDS, ANGELA
; TITLE OF INVENTION: CALCITONIN-RELATED MOLECULES
; FILE REFERENCE: A-684
; CURRENT APPLICATION NUMBER: US/09/847,712
; CURRENT FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: 60/201,511
; PRIOR FILING DATE: 2000-05-03
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-847-712-2

Query Match          92.8%; Score 1238; DB 10; Length 228;
Best Local Similarity 100.0%; Pred. No. 4e-33;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
DB 1 MDKTHCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
QY 61 DGVEVHNAKTKPREEQNTSYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKA 120
DB 61 DGVEVHNAKTKPREEQNTSYRVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKA 120
QY 121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 180
DB 121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 180
QY 181 SDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNYHTOKSLSLSPGK 228
DB 181 SDGSFFLYSKLTVDKSRWQQGNVFCSCVMHEALHNYHTOKSLSLSPGK 228

RESULT 10
US-09-840-277-109
; Sequence 109, Application US/09840277
; Patent No. US20020168363A1
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: KOHNO, TADAHIKO
; APPLICANT: LACEY, DAVID LEE
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: INTEGRIN/ADHESION ANTAGONISTS
; FILE REFERENCE: A-688A
; CURRENT APPLICATION NUMBER: US/09/840,277
; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: 60/198,919
; PRIOR FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-843-221A-2

Query Match          92.8%; Score 1238; DB 9; Length 228;
Best Local Similarity 100.0%; Pred. No. 4e-33;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
DB 1 MDKTHCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60

; Sequence 2, Application US/09843221A
; Publication No. US20030039654A1
; GENERAL INFORMATION:
; APPLICANT: KOSTENUK, PAUL
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: LACEY, DAVID LEE
; TITLE OF INVENTION: MODULATORS OF RECEPTORS FOR PARATHYROID HORMONE AND PARATHYROID H
; FILE REFERENCE: A-665B
; CURRENT APPLICATION NUMBER: US/09/843,221A
; CURRENT FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 60/266,673
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/214,860
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/200,053
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-843-221A-2

Query Match          92.8%; Score 1238; DB 9; Length 228;
Best Local Similarity 100.0%; Pred. No. 4e-33;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
DB 1 MDKTHCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYV 60
```

```
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; LENGTH: 739
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFG1 heavy chain - DNase I fusion
US-09-825-012-61

Query Match          93.3%; Score 1245; DB 10; Length 739;
Best Local Similarity 49.1%; Pred. No. 3.6e-32;
Matches 234; Conservative 6; Mismatches 4; Indels 233; Gaps 3;

Qy  2 DKTHCTPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYD 61
Db  241 DKTHCTPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYD 300
Qy  62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
Db  301 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 360
Qy  122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 181
Db  361 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 420
Qy  182 DGSFFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNHYTQKSLSLSPGKDWLK--AF----- 234
Db  421 DGSFFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNHYTQKSLSLSPGSGGLKIAAFNIQTF 480
Qy  235 -----YDKV----- 234
Db  481 GETKMSNATLVSYIVQILSRVDIALVQEVDRSHLTAVGKLLDNLNQADPTHYVYVSEPL 540
Qy  235 ----- 234
Db  541 GRNSYKERYLFVYRPDQVSAVDSYYDDGCEPCGNDTFNREPAIVRFESRTEVREFAIV 600
Qy  235 ----- 234
Db  601 PLHAAPGDAVAEIDALYDVYLDVQEKWGLEDMVLMGDFNAGCSYVRPSQWSSIRLWTSPT 660
Qy  235 -----YDKV----- 234
Db  661 FQWLIPDSADTTPPTHCAYDRIVVAGMILLRGAVVPDSALPFPNFAAYGLSDQLAQ 717

RESULT 5
US-09-840-277-2
; Sequence 2, Application US/09840277
; Patent No. US20020168363A1
; GENERAL INFORMATION:
; APPLICANT: FEIGE, ULRICH
; APPLICANT: KOHNO, TADAHIKO
; APPLICANT: LACEY, DAVID LEE
; APPLICANT: BOONE, THOMAS CHARLES
; TITLE OF INVENTION: INTEGRIN/ADHESION ANTAGONISTS
; FILE REFERENCE: A-689A
; CURRENT APPLICATION NUMBER: US/09/840,277
; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: 60/198,919
; PRIOR FILING DATE: 2000-04-21
; PRIOR APPLICATION NUMBER: 60/201,394
; PRIOR FILING DATE: 2000-05-03
```

```
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-840-277-2

Query Match          92.8%; Score 1238; DB 9; Length 228;
Best Local Similarity 100.0%; Pred. No. 4e-33;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  1 MDKTHCTPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYV 60
Db  1 MDKTHCTPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYV 60
Qy  61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
Db  61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
Qy  121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 180
Db  121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 180
Qy  181 SDGSFFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNHYTQKSLSLSPGK 228
Db  181 SDGSFFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNHYTQKSLSLSPGK 228

RESULT 6
US-09-847-249A-2
; Sequence 2, Application US/09847249A
; Publication No. US20030032588A1
; GENERAL INFORMATION:
; APPLICANT: MARSHALL, WILLIAM S.
; APPLICANT: STARK, KEVIN LEE
; TITLE OF INVENTION: GLUCAGON ANTAGONIST
; FILE REFERENCE: A-693
; CURRENT APPLICATION NUMBER: US/09/847,249A
; CURRENT FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: 60/201,436
; PRIOR FILING DATE: 2000-05-03
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-847-249A-2

Query Match          92.8%; Score 1238; DB 9; Length 228;
Best Local Similarity 100.0%; Pred. No. 4e-33;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  1 MDKTHCTPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYV 60
Db  1 MDKTHCTPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYV 60
Qy  61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
Db  61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
Qy  121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 180
Db  121 KGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 180
Qy  181 SDGSFFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNHYTQKSLSLSPGK 228
Db  181 SDGSFFLYSKLTVDKSRWQOGNVFSCSVMHAEALHNHYTQKSLSLSPGK 228

RESULT 7
US-09-840-669B-2
```


Db 421 DGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTQKSLSLSPCKGSGGLKIAAFNIQ 480
QY 229 -----KAFYDKV-----AEKLKEA 245
Db 481 TFGETKMSNATLVSYIVQILSRVDIALVOEVRDLSHTAVGKLLDNLNQADPTDTHYVYVSE 540
QY 229 -----KAFYDKV-----AEKLKEA 245
Db 541 PLGRNSYKERYLFVYRPDQVSAVDYDDGCEPCGNDTFNREPAIVRFFSRFTEVREFA 600
QY 229 -----KAFYDKV-----AEKLKEA 245
Db 601 IVP LHAAPGDAVAEIDALYDVLDVQEKWGLDVMLMGDFNAGCCSYVRPSSQSSIRLWTS 660
QY 229 -----KAFYDKV-----AEKLKEA 245
Db 661 PTFQWLIPDSADTTATPTHTCAIDRIVVAGMLLRGAVVPDSALPFFNQAAAYGLSDQLAQ 719
QY 229 -----KAFYDKV-----AEKLKEA 245
RESULT 2
US-09-825-012-55
; Sequence 55, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55
; LENGTH: 741
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFGL heavy chain - DNase I fusion
US-09-825-012-55
Query Match 93.4%; Score 1246; DB 10; Length 741;
Best Local Similarity 48.6%; Pred. No. 3.4e-32;
Matches 233; Conservative 6; Mismatches 5; Indels 235; Gaps 3;
QY 2 DKTHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61
Db 241 DKTHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 300
QY 62 GVEVHNATKPREEQNSTYRVSVLTVLHODWLNCKEYKCKVSNKALPAPIETKISKAK 121
Db 301 GVEVHNATKPREEQNSTYRVSVLTVLHODWLNCKEYKCKVSNKALPAPIETKISKAK 360
QY 122 GPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGOPENNYKTTTPVLD 181
Db 361 GPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGOPENNYKTTTPVLD 420
QY 182 DGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTQKSLSLSPCKGSGGLKIAAFNIQ 480
Db 421 DGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTQKSLSLSPCKGSGGLKIAAFNIQ 540
QY 229 -----KAFYDKV-----AEKLKEA 245
Db 481 TFGETKMSNATLVSYIVQILSRVDIALVOEVRDLSHTAVGKLLDNLNQADPTDTHYVYVSE 540
QY 229 -----KAFYDKV-----AEKLKEA 245
Db 541 PLGRNSYKERYLFVYRPDQVSAVDYDDGCEPCGNDTFNREPAIVRFFSRFTEVREFA 600
QY 229 -----KAFYDKV-----AEKLKEA 245

Db 601 IVP LHAAPGDAVAEIDALYDVLDVQEKWGLDVMLMGDFNAGCCSYVRPSSQSSIRLWTS 660
QY 229 -----KAFYDKV-----AEKLKEA 245
Db 661 PTFQWLIPDSADTTATPTHTCAIDRIVVAGMLLRGAVVPDSALPFFNQAAAYGLSDQLAQ 719
QY 229 -----KAFYDKV-----AEKLKEA 245
RESULT 3
US-09-825-012-52
; Sequence 52, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFGL heavy chain - DNase I fusion
US-09-825-012-52
Query Match 93.3%; Score 1245; DB 10; Length 729;
Best Local Similarity 49.1%; Pred. No. 3.5e-32;
Matches 234; Conservative 6; Mismatches 4; Indels 233; Gaps 3;
QY 2 DKTHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61
Db 241 DKTHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 300
QY 62 GVEVHNATKPREEQNSTYRVSVLTVLHODWLNCKEYKCKVSNKALPAPIETKISKAK 121
Db 301 GVEVHNATKPREEQNSTYRVSVLTVLHODWLNCKEYKCKVSNKALPAPIETKISKAK 360
QY 122 GPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGOPENNYKTTTPVLD 181
Db 361 GPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVWESNGOPENNYKTTTPVLD 420
QY 182 DGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTQKSLSLSPCKGSGGLKIAAFNIQ 480
Db 421 DGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTQKSLSLSPCKGSGGLKIAAFNIQ 540
QY 235 -----YDKV-----AEKLKEA 245
Db 481 GETKMSNATLVSYIVQILSRVDIALVOEVRDLSHTAVGKLLDNLNQADPTDTHYVYVSE 540
QY 235 -----YDKV-----AEKLKEA 245
Db 541 GRNSYKERYLFVYRPDQVSAVDYDDGCEPCGNDTFNREPAIVRFFSRFTEVREFA 600
QY 235 -----YDKV-----AEKLKEA 245
Db 601 PLHAAPGDAVAEIDALYDVLDVQEKWGLDVMLMGDFNAGCCSYVRPSSQSSIRLWTSPT 660
QY 235 -----YDKV-----AEKLKEA 245
Db 661 FQWLIPDSADTTATPTHTCAIDRIVVAGMLLRGAVVPDSALPFFNQAAAYGLSDQLAQ 717
QY 235 -----YDKV-----AEKLKEA 245
RESULT 4
US-09-825-012-61
; Sequence 61, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:43:21 ; Search time 11.3353 seconds
(without alignments)
1640.982 Million cell updates/sec

Title: 2LINK7
Perfect score: 1334
Sequence: 1 MDKTHCTPCPAPELLGGPS.....GKDLKAFYDKVAEXLKEAF 246

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.0

Searched: 288829 seqs, 75613885 residues
Total number of hits satisfying chosen parameters: 288829

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published_Applications_AA:*
- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1246	93.4	731	10	US-09-825-012-46
2	1246	93.4	741	10	US-09-825-012-55
3	1245	93.3	729	10	US-09-825-012-52
4	1245	93.3	739	10	US-09-825-012-61
5	1238	92.8	228	9	US-09-840-277-2
6	1238	92.8	228	9	US-09-847-249A-2
7	1238	92.8	228	9	US-09-840-669B-2
8	1238	92.8	228	9	US-09-843-221A-2
9	1238	92.8	228	10	US-09-847-712-2
10	1238	92.8	282	9	US-09-840-277-109
11	1238	92.8	389	10	US-09-784-623-14
12	1237	92.7	394	10	US-09-854-864-31
13	1237	92.7	730	10	US-09-825-012-49
14	1237	92.7	740	10	US-09-825-012-58
15	1234	92.5	229	9	US-10-215-297-2
16	1234	92.5	229	9	US-10-215-298-2
17	1234	92.5	347	9	US-10-091-236-17
18	1234	92.5	347	9	US-10-091-313-7
19	1234	92.5	347	9	US-10-091-268-7

ALIGNMENTS

RESULT 1
US-09-825-012-46
; Sequence 46, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 46
; LENGTH: 731
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFgl heavy chain - DNase I fusion
US-09-825-012-46
Query Match 93.4%; Score 1246; DB 10; Length 731;
Best Local Similarity 48.6%; Pred. No. 3 3e-32;
Matches 233; Conservative 6; Mismatches 5; Indels 235; Gaps 3;
QY 2 DKTHCTPCPAPELLGGPSVFLFPKPKDPTLMSRTPETVCVVVDVSHEDPEVKFNMYVD 61
Db 241 DKTHCTPCPAPELLGGPSVFLFPKPKDPTLMSRTPETVCVVVDVSHEDPEVKFNMYVD 300
QY 62 GVEVHNAKTPREQNSTYRVVSVLTVLDHDLWGKEYCKYKSNKALPAPIETISKAK 121
Db 301 GVEVHNAKTPREQNSTYRVVSVLTVLDHDLWGKEYCKYKSNKALPAPIETISKAK 360
QY 122 GQPREQVYTLPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDLS 181
Db 361 GQPREQVYTLPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDLS 420
QY 182 DGSFFLYSKLTVDKSRWQQGNFSCSVMHEALHNHYTQKSLSLSPGK----- 228
|||||

		Matches	227;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	2	DKTHTCPCPAP	ELGGPSVFLFP	PKPKD	TLMI	SRTPE	VTCTV	VDVDS	HEDPE	VKNWYVD	61
Db	150	DKTHTCPCPAP	ELGGPSVFLFP	PKPKD	TLMI	SRTPE	VTCTV	VDVDS	HEDPE	VKNWYVD	209
Qy	62	GVEVHNAKTK	PREQYNSTYR	VSVLTV	LHQD	WLNG	KEYCK	KVSN	KALPA	PIEKTISKAK	121
Db	210	GVEVHNAKTK	PREQYNSTYR	VSVLTV	LHQD	WLNG	KEYCK	KVSN	KALPA	PIEKTISKAK	269
Qy	122	GOPREPQVY	TLPPSRDEL	TKNQV	SLTCL	VLKGF	YPSD	IAV	WESNG	QPENNYK	181
Db	270	GOPREPQVY	TLPPSRDEL	TKNQV	SLTCL	VLKGF	YPSD	IAV	WESNG	QPENNYK	329
Qy	182	DGSFFLYSK	LTVDKSR	WOOGNV	FSCSV	MHEAL	HNHYT	OKSL	SLSPGK	228	
Db	330	DGSFFLYSK	LTVDKSR	WOOGNV	FSCSV	MHEAL	HNHYT	OKSL	SLSPGK	376	

Search completed: April 21, 2003, 10:50:21
Job time : 14.7471 secs

APPLICATION NUMBER: 07/936190
FILING DATE: 26-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/842777
FILING DATE: 18-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/250785
FILING DATE: 28-SEP-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/104329
FILING DATE: 02-OCT-1987
ATTORNEY/AGENT INFORMATION:
NAME: Hasak, Janet E.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 444P1C2
TELEPHONE: 415/225-1896
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 371 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-236-311-7

Query Match 92.4%; Score 1233; DB 1; Length 371;
Best Local Similarity 100.0%; Pred. No. 1.4e-42;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DKHTCPCPAPPELLGGPSVFLFPPPKPDTLMISRTPEVTCVVDVSHEDPEVKFNWYD 61
|||||
Db 145 DKHTCPCPAPPELLGGPSVFLFPPPKPDTLMISRTPEVTCVVDVSHEDPEVKFNWYD 204
|||||

QY 62 GVEVHNAKTPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
|||||
Db 205 GVEVHNAKTPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 264
|||||

QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 181
|||||
Db 265 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 324
|||||

QY 182 DGSFFLYSKLTVDKSRWQOGNPFSCSMHEALHNHYTQKSLSLSPGK 228
|||||
Db 325 DGSFFLYSKLTVDKSRWQOGNPFSCSMHEALHNHYTQKSLSLSPGK 371
|||||

RESULT 14
US-08-457-918-7
Sequence 7, Application US/08457918
Patent No. 6117655
GENERAL INFORMATION:
APPLICANT: Capon, Daniel J.
APPLICANT: Gregory, Timothy J.
TITLE OF INVENTION: Adhesion Variants
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,918
FILING DATE: 1-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/236311
FILING DATE: 02-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/936190
FILING DATE: 26-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/842777
FILING DATE: 18-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/250785
FILING DATE: 28-SEP-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/104329
FILING DATE: 02-OCT-1987
ATTORNEY/AGENT INFORMATION:
NAME: Kudinec, Jeffrey S.
REGISTRATION NUMBER: 36,575
REFERENCE/DOCKET NUMBER: P0444P1C3
TELEPHONE: 415/225-8228
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 371 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-457-918-7

Query Match 92.4%; Score 1233; DB 3; Length 371;
Best Local Similarity 100.0%; Pred. No. 1.4e-42;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DKHTCPCPAPPELLGGPSVFLFPPPKPDTLMISRTPEVTCVVDVSHEDPEVKFNWYD 61
|||||
Db 145 DKHTCPCPAPPELLGGPSVFLFPPPKPDTLMISRTPEVTCVVDVSHEDPEVKFNWYD 204
|||||

QY 62 GVEVHNAKTPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
|||||
Db 205 GVEVHNAKTPREEQNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 264
|||||

QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 181
|||||
Db 265 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 324
|||||

QY 182 DGSFFLYSKLTVDKSRWQOGNPFSCSMHEALHNHYTQKSLSLSPGK 228
|||||
Db 325 DGSFFLYSKLTVDKSRWQOGNPFSCSMHEALHNHYTQKSLSLSPGK 371
|||||

RESULT 15
US-09-180-100-22
Sequence 22, Application US/09180100
Patent No. 6306395
GENERAL INFORMATION:
APPLICANT: NAKAMURA, No. 630639510
APPLICANT: NAGATA, Shigekazu
TITLE OF INVENTION: NOVEL Fas ANTIGEN DERIVATIVE
FILE REFERENCE: 1110-207P
CURRENT APPLICATION NUMBER: US/09/180,100
CURRENT FILING DATE: 1998-11-02
EARLIER APPLICATION NUMBER: PCT/JP97/01502
EARLIER FILING DATE: 1997-05-01
NUMBER OF SEQ ID NOS: 25
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 22
LENGTH: 376
TYPE: PRT
ORGANISM: Homo sapiens
US-09-180-100-22

Query Match 92.4%; Score 1233; DB 4; Length 376;
Best Local Similarity 100.0%; Pred. No. 1.4e-42;

; CURRENT APPLICATION NUMBER: US/09/131,247
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: 60/055,185
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: PCT/US 97/02131
; EARLIER FILING DATE: 1997-02-10
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Human
US-09-131-247-6

Query Match 92.4%; Score 1233; DB 4; Length 235;
Best Local Similarity 100.0%; Pred. No. 4.3e-43;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 DKHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKENWYVD 61
DB 9 DKHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKENWYVD 68
QY 62 GVEVHNAKTPREEQNSTYRVSVTLVHQLDNLGKEYCKVSNKALPAPIEKTISKAK 121
DB 69 GVEVHNAKTPREEQNSTYRVSVTLVHQLDNLGKEYCKVSNKALPAPIEKTISKAK 128
QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 181
DB 129 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 188
QY 182 DGSFELYSLKTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSPGK 228
DB 189 DGSFELYSLKTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSPGK 235

RESULT 11
US-09-178-869-2
; Sequence 2, Application US/09178869B
; Patent No. 6197294
; GENERAL INFORMATION:
; APPLICANT: Tao, Weng
; APPLICANT: Wong, Shou
; APPLICANT: Hickey, William F.
; APPLICANT: Hamming, Joseph P.
; APPLICANT: Baetge, E. Edward
; TITLE OF INVENTION: CELL SURFACE-INDUCED MACROPHAGE ACTIVATION
; FILE REFERENCE: 17810-043
; CURRENT APPLICATION NUMBER: US/09/178,869B
; CURRENT FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-178-869-2

Query Match 92.4%; Score 1233; DB 4; Length 331;
Best Local Similarity 100.0%; Pred. No. 1e-42;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 DKHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKENWYVD 61
DB 105 DKHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKENWYVD 164
QY 62 GVEVHNAKTPREEQNSTYRVSVTLVHQLDNLGKEYCKVSNKALPAPIEKTISKAK 121
DB 165 GVEVHNAKTPREEQNSTYRVSVTLVHQLDNLGKEYCKVSNKALPAPIEKTISKAK 224
QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 181
DB 225 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 284

QY 182 DGSFELYSLKTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSPGK 228
DB 285 DGSFELYSLKTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSPGK 331
RESULT 12
US-09-180-100-11
; Sequence 11, Application US/09180100
; Patent No. 6306395
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, No. 6306395io
; APPLICANT: NAGATA, Shigekazu
; TITLE OF INVENTION: NOVEL Fas ANTIGEN DERIVATIVE
; FILE REFERENCE: 1110-207P
; CURRENT APPLICATION NUMBER: US/09/180,100
; CURRENT FILING DATE: 1998-11-02
; EARLIER APPLICATION NUMBER: PCT/JP97/01502
; EARLIER FILING DATE: 1997-05-01
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-180-100-11

Query Match 92.4%; Score 1233; DB 4; Length 360;
Best Local Similarity 100.0%; Pred. No. 1.3e-42;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 DKHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKENWYVD 61
DB 134 DKHTCCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKENWYVD 193
QY 62 GVEVHNAKTPREEQNSTYRVSVTLVHQLDNLGKEYCKVSNKALPAPIEKTISKAK 121
DB 194 GVEVHNAKTPREEQNSTYRVSVTLVHQLDNLGKEYCKVSNKALPAPIEKTISKAK 253
QY 122 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 181
DB 254 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVWESNGQPENNYKTTPPVLD 313
QY 182 DGSFELYSLKTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSPGK 228
DB 314 DGSFELYSLKTVDKSRWQGNVFCSCVMHEALHNYHTOKSLSPGK 360

RESULT 13
US-08-236-311-7
; Sequence 7, Application US/08236311
; Patent No. 5565335
; GENERAL INFORMATION:
; APPLICANT: Capon, Daniel J.
; APPLICANT: Gregory, Timothy J.
; TITLE OF INVENTION: Adhesion Variants
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/236,311
; FILING DATE: 02-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

Db 300 SDGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKSLSPGK 347

RESULT 8
US-09-189-129-2
; Sequence 2, Application US/09189129
; Patent No. 6323027
; GENERAL INFORMATION:
; APPLICANT: Burkly, Linda C
; APPLICANT: Benjamin, Christopher D
; APPLICANT: Hession, Catherine A
; APPLICANT: Whitty, Adrian
; TITLE OF INVENTION: COMMON GAMMA CHAIN BLOCKING AGENTS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Biogen, Inc.
; STREET: 14 Cambridge Center
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02142
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/189,129
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: A006 PCT CIP
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: 60/017,466
; FILING DATE: 10-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kaplan, Warren A.
; REGISTRATION NUMBER: 34,199
; REFERENCE/DOCKET NUMBER: A006 PCT CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 679-2838
; TELEFAX: 617 679-2838
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 482 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: YES
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
US-09-189-129-2

Query Match 92.5%; Score 1234; DB 4; Length 482;
Best Local Similarity 99.6%; Pred. No. 2.5e-42;
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKTHTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60
Db 255 VDKTHTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 314
QY 61 DGEVHNATKPREEOYNSTYRVSVLTVLHQDWLNGKEYCKVSNKALPAPIETKISKAK 120
Db 315 DGEVHNATKPREEOYNSTYRVSVLTVLHQDWLNGKEYCKVSNKALPAPIETKISKAK 374
QY 121 KGOPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 180
Db 375 KGOPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 434
QY 181 SDGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKSLSPGK 228
Db 435 SDGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKSLSPGK 482

RESULT 9
US-08-595-043A-50
; Sequence 50, Application US/08595043A
; Patent No. 5935824
; GENERAL INFORMATION:
; APPLICANT: SCARLATO, GREGORY D.
; TITLE OF INVENTION: PROTEIN EXPRESSION SYSTEM
; NUMBER OF SEQUENCES: 90
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/595,043A
; FILING DATE: 31-JAN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: SGAR-00371
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 232 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-595-043A-50
Query Match 92.4%; Score 1233; DB 2; Length 232;
Best Local Similarity 100.0%; Pred. No. 4.2e-43;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 DKTHTCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61
Db 6 DKTHTCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 65
QY 62 GVEVHNATKPREEOYNSTYRVSVLTVLHQDWLNGKEYCKVSNKALPAPIETKISKAK 121
Db 66 GVEVHNATKPREEOYNSTYRVSVLTVLHQDWLNGKEYCKVSNKALPAPIETKISKAK 125
QY 122 GGPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 181
Db 126 GGPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 185
QY 182 DGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKSLSPGK 228
Db 186 DGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKSLSPGK 232
RESULT 10
US-09-131-247-6
; Sequence 6, Application US/09131247
; Patent No. 6294170
; GENERAL INFORMATION:
; APPLICANT: Boone, Thomas C.
; APPLICANT: Hershenson, Susan
; APPLICANT: Bevilacqua, Michael P.
; APPLICANT: Collins, David S.
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR TREATING INFLAMMATORY
; DISEASES
; FILE REFERENCE: A-365F

QY 121 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 180
DB 240 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 299
QY 181 SDGSFFLYSLKLTVDKSRWQGNVFCSCVMHEALHNNHYTKQSLSLSPGK 228
DB 300 SDGSFFLYSLKLTVDKSRWQGNVFCSCVMHEALHNNHYTKQSLSLSPGK 347

RESULT 6

US-08-466-465-8
; Sequence 8, Application US/08466465
; Patent No. 6162432
; GENERAL INFORMATION:
; APPLICANT: Wallner, Barbara P.
; APPLICANT: Cooper, Kevin D.
; TITLE OF INVENTION: Method of Prophylaxis or Treatment of Antigen
; TITLE OF INVENTION: Presenting Cell Driven Skin Conditions Using
; TITLE OF INVENTION: Inhibitors of the CD2/LFA-3 Interaction
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466.465
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08755
; FILING DATE: 06-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/862.022
; FILING DATE: 12-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770.969
; FILING DATE: 07-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis (PLM)
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGP-111CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-466-465-8

Query Match 92.5%; Score 1234; DB 4; Length 347;
Best Local Similarity 99.6%; Pred. No. 1.le-42;
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHTCPPCAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60
DB 120 VDKTHTCPPCAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 179
QY 61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIETKISKA 120
DB 180 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIETKISKA 239
QY 121 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 180

DB 240 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 299
QY 181 SDGSFFLYSLKLTVDKSRWQGNVFCSCVMHEALHNNHYTKQSLSLSPGK 228
DB 300 SDGSFFLYSLKLTVDKSRWQGNVFCSCVMHEALHNNHYTKQSLSLSPGK 347

RESULT 7

PCT-US92-02050-43
; Sequence 43, Application PC/TUS9202050
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: WALLNER, Barbara P.
; APPLICANT: MILLER, Glenn T.
; APPLICANT: ROSA, Margaret D.
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
; TITLE OF INVENTION: FUNCTION ASSOCIATED ANTIGEN 3
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 875 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022-6250
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02050
; FILING DATE: 19920312
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/667.971
; FILING DATE: 12-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770.967
; FILING DATE: 07-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY, James F., Jr.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B151CIP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)715-0600
; TELEFAX: (212)715-0673
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US92-02050-43

Query Match 92.5%; Score 1234; DB 5; Length 347;
Best Local Similarity 99.6%; Pred. No. 1.le-42;
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDKTHTCPPCAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60
DB 120 VDKTHTCPPCAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 179
QY 61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIETKISKA 120
DB 180 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIETKISKA 239
QY 121 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 180
DB 240 KGQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPPVLD 299
QY 181 SDGSFFLYSLKLTVDKSRWQGNVFCSCVMHEALHNNHYTKQSLSLSPGK 228

QY 239 -----AEKLKE 244
DB 310 GDETRQLQLEAVNITDLSENKQ 331

RESULT 2
US-07-940-861-43
; Sequence 43, Application US/07940861
; Patent No. 5547853
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: WALLNER, Barbara P.
; APPLICANT: MILLER, Glenn T.
; APPLICANT: ROSA, Margaret D.
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 875 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022-6250
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/940,861
; FILING DATE: 21-OCT-1992
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02050
; FILING DATE: 12-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/667,971
; FILING DATE: 12-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,967
; FILING DATE: 07-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY, James F., Jr.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B151CIP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)715-0600
; TELEFAX: (212)715-0673
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-940-861-43

Query Match 92.5%; Score 1234; DB 1; Length 347;
Best Local Similarity 99.6%; Pred. No. 1.1e-42;
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWTV 60
DB 120 VDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWTV 179
QY 61 DGVEVHNAKTKPREEQNTSYRYVSVLTVLHQDNLNGKEYCKVSNKALPAPIEKTISKA 120
DB 180 DGVEVHNAKTKPREEQNTSYRYVSVLTVLHQDNLNGKEYCKVSNKALPAPIEKTISKA 239
QY 121 KGQPREPQVYTLPPSDELTKNQVSLTCLVKGFPSPDIKAVENESGQPNNNYKTTTPPVL 180
DB 240 KGQPREPQVYTLPPSDELTKNQVSLTCLVKGFPSPDIKAVENESGQPNNNYKTTTPPVL 299

QY 181 SDGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKSLSLSPGK 228
DB 300 SDGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKSLSLSPGK 347

RESULT 3
US-08-459-512-43
; Sequence 43, Application US/08459512
; Patent No. 5728677
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: WALLNER, Barbara P.
; APPLICANT: MILLER, Glenn T.
; APPLICANT: ROSA, Margaret D.
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Neave
; STREET: 875 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022-6250
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,512
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02050
; FILING DATE: 12-MAR-1992
; APPLICATION NUMBER: US 07/667,971
; FILING DATE: 12-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/770,967
; FILING DATE: 07-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: HALEY, James F., Jr.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B151CIP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)715-0600
; TELEFAX: (212)715-0673
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 347 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-459-512-43

Query Match 92.5%; Score 1234; DB 1; Length 347;
Best Local Similarity 99.6%; Pred. No. 1.1e-42;
Matches 227; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWTV 60
DB 120 VDKTHTCPCPAPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWTV 179
QY 61 DGVEVHNAKTKPREEQNTSYRYVSVLTVLHQDNLNGKEYCKVSNKALPAPIEKTISKA 120
DB 180 DGVEVHNAKTKPREEQNTSYRYVSVLTVLHQDNLNGKEYCKVSNKALPAPIEKTISKA 239
QY 121 KGQPREPQVYTLPPSDELTKNQVSLTCLVKGFPSPDIKAVENESGQPNNNYKTTTPPVL 180
DB 240 KGQPREPQVYTLPPSDELTKNQVSLTCLVKGFPSPDIKAVENESGQPNNNYKTTTPPVL 299

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 10:42:26 ; Search time 13.7471 Seconds
(without alignments)
526.515 Million cell updates/sec

Title: 2LINK7

Perfect score: 1334

Sequence: 1 MDKTHCPCPAPPELLGGPS.....GRDLKAFYDKVAELKEAF 246

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA:*

- 1: /cgn2.6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2.6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2.6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2.6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2.6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2.6/ptodata/1/iaa/backfiles1.pep.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1238	92.8	389	4 US-09-131-247-14	Sequence 14, Appl
2	1234	92.5	347	1 US-07-940-861-43	Sequence 43, Appl
3	1234	92.5	347	1 US-08-459-512-43	Sequence 43, Appl
4	1234	92.5	347	2 US-08-459-657-43	Sequence 43, Appl
5	1234	92.5	347	2 US-08-460-132-43	Sequence 43, Appl
6	1234	92.5	347	4 US-08-466-465-8	Sequence 8, Appl
7	1234	92.5	347	5 PCT-US92-02050-43	Sequence 43, Appl
8	1234	92.5	482	4 US-09-189-129-2	Sequence 2, Appl
9	1233	92.4	232	2 US-08-595-043A-50	Sequence 50, Appl
10	1233	92.4	235	4 US-09-131-247-6	Sequence 6, Appl
11	1233	92.4	331	4 US-09-178-869-2	Sequence 2, Appl
12	1233	92.4	360	4 US-09-180-100-11	Sequence 11, Appl
13	1233	92.4	371	1 US-08-236-311-7	Sequence 7, Appl
14	1233	92.4	371	3 US-08-457-918-7	Sequence 7, Appl
15	1233	92.4	376	4 US-09-180-100-22	Sequence 22, Appl
16	1233	92.4	387	1 US-08-470-299-4	Sequence 4, Appl
17	1233	92.4	388	4 US-09-131-247-16	Sequence 16, Appl
18	1233	92.4	396	2 US-08-784-512-3	Sequence 3, Appl
19	1233	92.4	396	4 US-09-176-228-3	Sequence 3, Appl
20	1233	92.4	424	4 US-09-333-593A-8	Sequence 8, Appl
21	1233	92.4	424	5 PCT-US95-03866-12	Sequence 12, Appl
22	1233	92.4	424	5 PCT-US95-03866-14	Sequence 14, Appl
23	1233	92.4	437	5 PCT-US96-10043-11	Sequence 11, Appl
24	1233	92.4	442	5 PCT-US96-10043-9	Sequence 9, Appl
25	1233	92.4	446	3 US-08-397-411-7	Sequence 7, Appl
26	1233	92.4	449	1 US-08-458-516-13	Sequence 13, Appl
27	1233	92.4	459	1 US-08-157-101A-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-09-131-247-14
; Sequence 14, Application US/09131247
; Patent No. 6294170
; GENERAL INFORMATION:
; APPLICANT: Boone, Thomas C.
; APPLICANT: Hersenson, Susan
; APPLICANT: Bevilacqua, Michael P.
; APPLICANT: Collins, David S.
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR TREATING INFLAMMATORY
; TITLE OF INVENTION: DISEASES
; FILE REFERENCE: A-365F
; CURRENT APPLICATION NUMBER: US/09/131.247
; EARLIER FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: 60/055,185
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: PCT/US 97/02131
; EARLIER FILING DATE: 1997-02-10
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Human
US-09-131-247-14

Query Match 92.8%; Score 1238; DB 4; Length 389;
Best Local Similarity 72.7%; Pred. No. 9.7e-43;
Matches 234; Conservative 4; Mismatches 5; Indels 79; Gaps 4;

QY	2	DKTHTCPCPAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVD	61
Db	10	DKTHTCPCPAPPELLGGPSVFLFPPPKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVD	69
QY	62	GVEVHNAKTPREEQNSTYRVVSVLTVLHQDMLNKGKCKVSNKALPAPIEKTISKAK	121
Db	70	GVEVHNAKTPREEQNSTYRVVSVLTVLHQDMLNKGKCKVSNKALPAPIEKTISKAK	129
QY	122	GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGQPENNYKTTTPPVLD	181
Db	130	GQPREQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEMESNGQPENNYKTTTPPVLD	189
QY	182	DGSEFFLYSKLTVDKSRWQGNVFSCVSMHEALHNHYTQKSLSLSPGK-----	228
Db	190	DGSEFFLYSKLTVDKSRWQGNVFSCVSMHEALHNHYTQKSLSLSPGKMRPSGRKSKMQA	249
QY	229	--DW---LKAFY-----DKV-----	238
Db	250	FRWDVYNQKTFYLRNNQLVAGYLGQPNVNLEEKIDVYVPIEPIHALFLGIHGKMKLSCVKS	309


```
QY 182 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSLSPGK----- 228
Db 209 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKSLSLSPGKTSEETISTVQEKQ 268
QY 229 -----DW----- 230
Db 269 QNISPVRERGPORVAAHITGTRGRSNTLSSPNSKNEKALGRKINSWESSRSGHSFLSNL 328
QY 231 -----LKAFY----- 235
Db 329 HLRNGELVIHEKGFYIYSQTYFRFOEEIKENTKNDQMVOYIYKYTSPDPILLMKSAR 388
QY 236 -----DKVAE-----KLKE 244
Db 389 NSCWSKDAEYGLYSIQGGIFELKE 413
```

Search completed: April 21, 2003, 10:45:18
Job time : 30.9412 secs

KW cancer; hepatitis; acquired immunodeficiency syndrome; AIDS;
 KW autoimmune disorder; transplant rejection; cardiovascular disease;
 KW arteriosclerosis; Fc-huAGP-1; fusion protein.

OS Homo sapiens.

XX WO2000063253-A1.

XX 26-OCT-2000.

XX 24-MAR-2000; 2000WO-US08004.

PR 16-APR-1999; 99US-0293245.

XX (AMGE-) AMGEN INC.

XX Hsu H, Meng S;

XX WPI; 2000-665240/64.

DR N-PSDB; AAC67833.

XX Fusion protein of AGP-1 protein and an Fc region, used to treat
 PT proliferative disorders, immune disorders, and virally-induced
 PT disorders -

XX Disclosure; Fig 4; 93pp; English.

XX The present sequence is an AGP-1 fusion protein. AGP-1 is a
 GC type II transmembrane protein. The fusion proteins comprise an Fc
 GC immunoglobulin region fused to the N-terminal portion of the AGP-1
 GC protein. The fusion proteins can be used to induce apoptosis in a tissue,
 GC and to treat proliferative disorders, immune disorders, or
 GC virally-induced disorders. The proliferative disorders include cancers,
 GC such as breast, prostate, lung or colon cancer. The viral infections
 GC include hepatitis, and acquired immunodeficiency syndrome (AIDS), and the
 GC immune disorders may be autoimmune disorders or transplant rejection.
 CC Cardiovascular diseases such as arteriosclerosis may also be treated. The
 CC AGP-1 containing fusion proteins have increased biological activity
 CC compared to the soluble AGP-1 proteins used in prior art therapies.

XX Sequence 423 AA;

Query Match 93.3%; Score 1245; DB 21; Length 423;
 Best Local Similarity 64.8%; Pred. No. 6e-35;
 Matches 237; Conservative 1; Mismatches 5; Indels 123; Gaps 4;

Qy 2 DKHTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61

Db 29 DKHTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 88

Qy 62 GVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIETISKAK 121

Db 89 GVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIETISKAK 148

Qy 122 GQPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 181

Db 149 GQPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 208

Qy 182 DGSFELYSLTKVDRKRWQGNVFCVSMHEALHNYHTOKSLSPGK----- 228

Db 209 DGSFELYSLTKVDRKRWQGNVFCVSMHEALHNYHTOKSLSPGKVRERQPVAAHI 268

Qy 229 -----DW-----LKAFY----- 235

Db 269 TGTRGRSNTLSPNSKNEKALGRKINSWESSRSGHSFLSLNHLRNGELVIHEKGYIYS 328

Qy 236 -----DKVAE----- 240

Db 329 QTYFRFQBEIKENTKNDKQVQYIYKTYSPDPILLMKRSARNCSWCKDAEYGLYSIQGG 388

Qy 241 --KLKE 244

Db 389 IFELKE 394

RESULT 15

AAB28692

ID AAB28692 standard; Protein; 441 AA.

XX AC AAB28692;

XX 14-FEB-2001 (first entry)

DE FC-huAGP-1 (95-281) fusion protein.

XX Human; AGP-1; type II transmembrane protein; cytostatic; antiviral;
 KW antiinflammatory; hepatotropic; antiarteriosclerotic; anti-HIV; HIV;
 KW human immunodeficiency virus; apoptosis; proliferative disorder;
 KW cancer; hepatitis; acquired immunodeficiency syndrome; AIDS;
 KW autoimmune disorder; transplant rejection; cardiovascular disease;
 KW arteriosclerosis; Fc-huAGP-1; fusion protein.

XX Homo sapiens.

XX WO2000063253-A1.

XX 26-OCT-2000.

XX 24-MAR-2000; 2000WO-US08004.

XX 16-APR-1999; 99US-0293245.

XX (AMGE-) AMGEN INC.

XX Hsu H, Meng S;

XX WPI; 2000-665240/64.

DR N-PSDB; AAC67832.

XX Fusion protein of AGP-1 protein and an Fc region, used to treat
 PT proliferative disorders, immune disorders, and virally-induced
 PT disorders -

XX Disclosure; Fig 3; 93pp; English.

XX The present sequence is an AGP-1 fusion protein. AGP-1 is a
 CC type II transmembrane protein. The fusion proteins comprise an Fc
 CC immunoglobulin region fused to the N-terminal portion of the AGP-1
 CC protein. The fusion proteins can be used to induce apoptosis in a tissue,
 CC and to treat proliferative disorders, immune disorders, or
 CC virally-induced disorders. The proliferative disorders include cancers,
 CC such as breast, prostate, lung or colon cancer. The viral infections
 CC include hepatitis, and acquired immunodeficiency syndrome (AIDS), and the
 CC immune disorders may be autoimmune disorders or transplant rejection.
 CC Cardiovascular diseases such as arteriosclerosis may also be treated. The
 CC AGP-1 containing fusion proteins have increased biological activity
 CC compared to the soluble AGP-1 proteins used in prior art therapies.

XX Sequence 441 AA;

Query Match 93.3%; Score 1245; DB 21; Length 441;
 Best Local Similarity 61.6%; Pred. No. 6.6e-35;
 Matches 237; Conservative 1; Mismatches 5; Indels 142; Gaps 4;

Qy 2 DKHTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 61

Db 29 DKHTCCPCAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYVD 88

Qy 62 GVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIETISKAK 121

Db 89 GVEVHNATKPREQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIETISKAK 148

Qy 122 GQPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 181

Db 149 GQPREPQVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD 208

Db 121 KGQPREPOVYTLPPSRDELTKNQSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 180
QY 181 SDGSFPLYSLTVDKSRWQGNVFCVSMHEALHNHYTOKSLSPGKD----- 229
Db 181 SDGSFPLYSLTVDKSRWQGNVFCVSMHEALHNHYTOKSLSPGKETFPKYLHYDE 240
QY 230 -----WLK----- 232
Db 241 ETSHQLLCKDPCPGTYLKQCTAKWTKVACPCPDHYTDSWHTSDCLCYCSPVKELQYV 300
QY 233 ----- 232
Db 301 KQECNRTHNRVCECKEGRYLEIEFCLKHRSCPPGFGVQAGTPERNTVCKRCPDGFFSNE 360
QY 233 -----AFYDKVAEKLKEA 245
Db 361 TSSKAPCRKHTNCSVFGLLLTKGNATHDNCISGNS 398
RESULT 13
AAB80904
ID AAB80904 standard; protein; 401 AA.
XX AAB80904;
XX
XX 31-MAY-2001 (first entry)
XX Human metFdeltaC-OPG(22-194) fusion protein.
XX
KW Human; anticancer; Antimetastatic; Osteogenic; lytic bone disease;
KW multiple myeloma; osteosclerotic bone metastasis; OPG; osteoprotegrin;
KW osteoclast formation inhibition; bone resorption inhibition;
KW immunoglobulin.
XX
OS Homo sapiens.
XX WO200117543-A2.
XX
XX 15-MAR-2001.
XX
XX 18-AUG-2000; 2000WO-US22806.
XX
XX 03-SEP-1999; 99US-0389545.
XX
XX (AMGE-) AMGEN INC.
XX
XX Dunstan CR;
XX
XX WPI; 2001-265936/27.
XX
XX Preventing or treating lytic bone diseases, particularly associated
XX with cancer or metastasis, by administering an osteoprotegrin
XX polypeptide -
XX
XX Claim 11; Fig 8; 87pp; English.
XX
XX The present invention relates to a method for the prevention or treatment
XX of lytic bone disease or multiple myeloma. Also the method can be used
XX for preventing metastasis of cancer to bone or osteosclerotic bone
XX metastasis. The method comprises administering an OPG (osteoprotegrin)
XX polypeptide or OPG fusion protein. The present sequence is one such OPG
XX fusion protein. OPG inhibits formation of osteoclasts (and thus bone
XX resorption) by blocking differentiation from monocytes/macrophage
XX precursors. The OPG polypeptide can be used in a method of preventing or
XX treating lytic bone disease, for preventing metastasis of cancer to bone
XX (e.g. breast, prostate, thyroid, cancer of the kidney, lung, oesophageal,
XX rectal, bladder, cervical, ovarian, liver, cancer of the gastrointestinal
XX tract, multiple myeloma or lymphoma) and preventing the osteosclerotic
XX bone metastasis. The OPG fusion polypeptides are used in the prevention
XX or treatment of loss of bone mass, which occurs in conditions including
XX osteoporosis, such as primary osteoporosis, endocrine osteoporosis
XX (hyperparathyroidism, Cushing's syndrome and acromegaly), hereditary and

congenital forms of osteoporosis (osteogenesis imperfecta,
homocystinuria, Menkes' syndrome and Riley-Day syndrome) and osteoporosis
due to immobilisation of extremities; Paget's disease of bone (osteitis
deformans) in adults and juveniles; osteomyelitis, or an infectious
lesion in bone; hypercalcaemia resulting from solid tumours (breast, lung
and kidney) and haematologic malignancies (multiple myeloma, lymphoma and
leukaemia), idiopathic hypercalcaemia, and hypercalcaemia associated with
hyperthyroidism and renal function disorders; osteopaenia associated with
surgery, induced by steroid administration, and associated with disorders
of the small and large intestine and with chronic hepatic and renal
diseases; osteonecrosis, or bone cell death, associated with traumatic
injury or nontraumatic necrosis associated with Gaucher's disease, sickle
cell anaemia, systemic lupus erythematosus and other conditions; bone
loss due to rheumatoid arthritis; periodontal bone loss; osteoarthritis;
prosthetic loosening; and osteolytic metastasis. The OPG fusion proteins
are also used in the replacement of structurally sound bone with
disorganised bone as seen in Paget's disease of bone (osteitis deformans)
in adults and juveniles; hyperparathyroidism, in congenital bone
disorders such as fibrous dysplasia, and in osteosclerotic bone
metastases. The OPG fusion proteins can exhibit increased circulating
half-lives and slower clearance times, thereby providing a more sustained
activity. The OPG fusion protein comprises a fragment of the human OPG
protein and the Fc region of immunoglobulin IgGgamma1 (the hinge, CH2 and
CH3 regions; see AAB80897-8).
XX
XX
SQ Sequence 401 AA;
Query Match 93.3%; Score 1245; DB 22; Length 401;
Best Local Similarity 58.5%; Pred. No. 5.3e-35;
Matches 233; Conservative 5; Mismatches 7; Indels 153; Gaps 2;
QY 1 MDKTHTCPAPPELGGPSVFLFPPKPKYDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60
Db 1 MDKTHTCPAPPELGGPSVFLFPPKPKYDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60
QY 61 DGVEVHNATKPREEQNSTYRVSVTLVTHQDLNKGKEYCKVSNKALPAPIEKTISKA 120
Db 61 DGVEVHNATKPREEQNSTYRVSVTLVTHQDLNKGKEYCKVSNKALPAPIEKTISKA 120
QY 121 KGQPREPOVYTLPPSRDELTKNQSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 180
Db 121 KGQPREPOVYTLPPSRDELTKNQSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD 180
QY 181 SDGSFPLYSLTVDKSRWQGNVFCVSMHEALHNHYTOKSLSPGKD----- 229
Db 181 SDGSFPLYSLTVDKSRWQGNVFCVSMHEALHNHYTOKSLSPGKETFPKYLHYDE 240
QY 230 -----WLK----- 232
Db 241 ETSHQLLCKDPCPGTYLKQCTAKWTKVACPCPDHYTDSWHTSDCLCYCSPVKELQYV 300
QY 233 ----- 232
Db 301 KQECNRTHNRVCECKEGRYLEIEFCLKHRSCPPGFGVQAGTPERNTVCKRCPDGFFSNE 360
QY 233 -----AFYDKVAEKLKEA 245
Db 361 TSSKAPCRKHTNCSVFGLLLTKGNATHDNCISGNS 398
RESULT 14
AAB28693
ID AAB28693 standard; Protein; 423 AA.
XX AAB28693;
XX
XX 14-FEB-2001 (first entry)
XX Fc-huAGP-1 (114-281) fusion protein.
XX Human; AGP-1; type II transmembrane protein; cytostatic; antiviral;
XX antinflammatory; hepatotropic; antiarteriosclerotic; anti-HIV; HIV;
XX human immunodeficiency virus; apoptosis; proliferative disorder;

```

OS Homo sapiens.
XX Synthetic.
XX Key
XX Location/Qualifiers
XX 2..374
XX Protein
XX /note= "Recombinant human Fc-OB variant 2 protein"
XX Region
XX 229..374
XX /note= "Human OB protein"
XX
XX W09828427-A1.
XX
XX 02-JUL-1998.
XX
XX 11-DEC-1997; 97WO-US23183.
XX
XX 20-DEC-1996; 96US-0770973.
XX
XX (AMGE-) AMGEN INC.
XX
XX Hecht RI, Mann MB;
XX
XX WPI: 1998-377658/32.
XX N-PSDB; AAV32902.
XX
XX New fusion proteins of OB and Fc - used for treating e.g. excess
XX weight, diabetes, arterial sclerosis, arterial plaque, high blood
XX lipid level, gall stones or stroke
XX
XX Claim 2; Fig 5A-5C; 107pp; English.
XX
XX The present sequence represents a recombinant human MetFc-OB variant 2
XX fusion protein having a 5 residue deletion of residues 2-6 of the
XX wild-type Fc-OB protein sequence shown in AAM49073. The invention
XX provides Fc-OB fusion proteins whereby the Fc region of an immunoglobulin
XX or its analogue is linked, either directly or indirectly using a linker,
XX to the N-terminus of an OB protein or its analogue. The Fc-OB fusion
XX proteins are claimed to demonstrate increased stability and clearance
XX rate and decreased degradation as compared to OB protein or a fusion of
XX Fc to the C-terminus of the OB protein. These Fc-OB fusion proteins are
XX also claimed to be useful for treating excess weight in an individual or
XX animal or for treating co-morbidities associated with excess fat such as
XX diabetes, high blood lipid level, arterial sclerosis and stroke.
XX
XX Sequence 374 AA;
XX
XX Query Match 93.3%; Score 1245; DB 19; Length 374;
XX Best Local Similarity 74.4%; Pred. No. 4.5e-35;
XX Matches 233; Conservative 4; Mismatches 7; Indels 59; Gaps 2;
XX
XX 1 MDKTHCTCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60
XX 1 MDKTHCTCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60
XX
XX 61 DGVEVHNATKPREEQNSTYRVSVSLTVLHODWLNKGEYCKVSNKALPAPIETISK 120
XX 61 DGVEVHNATKPREEQNSTYRVSVSLTVLHODWLNKGEYCKVSNKALPAPIETISK 120
XX
XX 121 KGQPREQVYTLPPSRDELTKNOVSLTCLVKGFPSPDI AVEWESNGQPNKYKTPPVLD 180
XX 121 KGQPREQVYTLPPSRDELTKNOVSLTCLVKGFPSPDI AVEWESNGQPNKYKTPPVLD 180
XX
XX 181 SDGSFLLYSKLTVDKSRWQGNVFCSCVHNEALHNYHTQKSLSPGK----- 228
XX 181 SDGSFLLYSKLTVDKSRWQGNVFCSCVHNEALHNYHTQKSLSPGKVPQKVDQDTKT 240
XX
XX 229 -----DWLKAFFDKV----- 238
XX
XX 241 LIKTIIVTRINDISHTQSVSSKOKVTGLDFIPGLHPIILTSKMDQTLAVYQQTSMPSKN 300
XX
XX 239 -----AEKLKE 244
XX 1:::
XX 301 VIQISNDLENLRD 313

```

```

RESULT 12
AAV72922
ID AAV72922 standard; Protein; 401 AA.
XX
XX AAV72922;
AC
XX
XX 13-JUN-2001 (first entry)
DT
XX
XX Human met-Fc (lacking 1-5 residues)-OPG (22-194 aa) fusion protein.
DE
XX
XX Human; fusion protein; osteoprotegerin; OPG; Fc protein; osteopathic;
XX therapy; bone loss; osteoporosis; Paget's disease; osteomyelitis;
XX hypercalcaemia; osteopenia; osteonecrosis; rheumatoid arthritis;
XX osteolytic metastasis; prosthetic loosening; immunoglobulin G1;
XX IgG1; periodontal.
XX
XX Homo sapiens.
OS
XX
XX Location/Qualifiers
XX Key
XX Region
XX /note= "Derived from human osteoprotegerin fragment
XX (22-194 residues)"
XX
XX Region
XX 1..228
XX /note= "Met-human IgG1 Fc region lacking 1-5 residues;
XX Corresponds to 6-231 residues of human IgG1 Fc region"
XX
XX W0200118203-A1.
XX
XX 15-MAR-2001.
PD
XX
XX 18-AUG-2000; 2000WO-US22797.
XX
XX 03-SEP-1999; 99US-0389782.
XX
XX (AMGE-) AMGEN INC.
XX
XX Dunstan CR, Wooden SK, Mann MB;
XX
XX WPI: 2001-244572/25.
XX
XX Osteoprotegerin-Fc protein fusions useful for treating bone loss caused
XX by e.g. osteoporosis, Paget's disease and osteomyelitis -
XX
XX Claim 7; Fig 8; 119pp; English.
XX
XX The present sequence is a fusion protein comprising human met-Fc region
XX (lacking 1-5 residues) which is fused with a sequence derived from
XX human osteoprotegerin (OPG; 22-194 residues) by a linker. OPG negatively
XX regulates the formation of osteoclasts in vitro and in vivo. It blocks
XX the differentiation of osteoclasts from monocyte or macrophage precursors
XX and the reabsorption of bone. The OPG-Fc fusion protein is administered
XX for the treatment of bone loss resulting from osteoporosis, Paget's
XX disease, osteomyelitis, hypercalcaemia, osteopenia associated with
XX surgery or steroid administration, osteonecrosis, bone loss due to
XX rheumatoid arthritis, periodontal bone loss, osteolytic metastasis
XX and/or prosthetic loosening.
XX
XX Sequence 401 AA;
XX
XX Query Match 93.3%; Score 1245; DB 22; Length 401;
XX Best Local Similarity 58.5%; Pred. No. 5.3e-35;
XX Matches 233; Conservative 5; Mismatches 7; Indels 153; Gaps 2;
XX
XX 1 MDKTHCTCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60
XX 1 MDKTHCTCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60
XX
XX 61 DGVEVHNATKPREEQNSTYRVSVSLTVLHODWLNKGEYCKVSNKALPAPIETISK 120
XX 61 DGVEVHNATKPREEQNSTYRVSVSLTVLHODWLNKGEYCKVSNKALPAPIETISK 120
XX
XX 121 KGQPREQVYTLPPSRDELTKNOVSLTCLVKGFPSPDI AVEWESNGQPNKYKTPPVLD 180
XX

```

CC cell-specific portion, comprising an humanised monoclonal antibody,
CC having specificity for polymorphic epithelial mucin (PEM) or its antigen
CC binding fragment and a cytotoxic portion having endonucleolytic activity,
CC exemplified by AAM52154-AM52168 and encoded by ABA02682-ABA02728. The
CC compound has cytostatic activity useful for treating cancer and acting as
CC a potential inducer of apoptosis.
XX
SQ Sequence 741 AA;

Query Match 93.4%; Score 1246; DB 22; Length 741;
Best Local Similarity 48.6%; Pred. No. 2e-34;
Matches 233; Conservative 6; Mismatches 5; Indels 235; Gaps 3;

QY 2 DKHTCTCPAPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKENWYVD 61
DB 241 DKHTCTCPAPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKENWYVD 300
QY 62 GVEVHNAKTPREEQNSTYRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKAK 121
DB 301 GVEVHNAKTPREEQNSTYRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKAK 360
QY 122 GQREPOVYTLPPSRDELTKNOVSLCLVKGFPSPDIWESNGQPNENYKTTTPPVLD 181
DB 361 GQREPOVYTLPPSRDELTKNOVSLCLVKGFPSPDIWESNGQPNENYKTTTPPVLD 420
QY 182 DGSFFLYSKLTVDKSRWQOGNVSFSCVMHEALHNHYTQKSLSLSPGK----- 228
DB 421 DGSFFLYSKLTVDKSRWQOGNVSFSCVMHEALHNHYTQKSLSLSPGKGLKIAAFNIQ 480
QY 229 ----- 228
DB 481 TFGETKMSNATLSYIVQLSRYDIALVQEVDRSHLTAVGKLLDNLNQADPTHYVWSE 540
QY 229 ----- 228
DB 541 PLGRNSYKERYLFVYRPDQVSAVDYDDGCEPCGNDTFNREPAIVRFSTFVREFA 600
QY 229 ----- 228
DB 601 IVPLHAAPGDAVAEIDALYDVLDOEKWGLEDMVLMGDFNACGSYVRPSOWSSIRLWTS 660
QY 229 ---DWL-----KAFYDKV-----AEKLKEA 245
DB 661 PTFQWLIPDSADTTATPTHCAYDIRIVAGMLLRGAVVPDSALPFPNQAAYGLSDQLAQA 719

RESULT 10
AAW83963
ID AAW83963 standard; Protein; 374 AA.
AC AAW83963;
XX
XX 05-FEB-1999 (first entry)
DE Recombinant human metFc-OB protein variant.
XX
KW Recombinant; metFc-OB protein; Fc region; immunoglobulin; Ig; OB;
KW obesity; human; adiposity; blood lipid; diabetes type II; insulin;
KW hypoglycaemic; antihypertensive; diuretic; appetite suppressant;
KW suspension; variant.
XX
OS Homo sapiens.
XX
PN WO9846257-A1.
XX
PD 22-OCT-1998.
XX
PF 16-APR-1998; 98WO-US07828.
XX
PR 14-APR-1998; 98US-0059467.
PR 17-APR-1997; 97US-0843971.
XX
PA (AMGE-) AMGEN INC.

XX Brems DN, French DL, Speed MA;
PI WPI; 1998-594525/50.
XX N-PSDB; AAV69686.
XX Concentrated suspension of fusion of obesity protein with Fc
PT immunoglobulin fragment - stable at physiological pH, used for e.g.
PT reduction of weight and blood lipid levels, and for treatment of
PT type II diabetes
XX Claim 2; Fig 6A-C; 47pp; English.
XX This represents a recombinant metFc-OB protein variant which consists of
CC an Fc region of human immunoglobulin (Ig) fused to a human OB (obesity)
CC protein. The invention provides a human OB protein suspension that
CC contains at least 0.5 mg/ml of the human OB protein derivatised by
CC attachment of the Fc region of an Ig to the N-terminus of OB, and has a
CC pH 6-8. The suspensions are used to reduce weight, adiposity and blood
CC lipid levels, to treat or prevent diabetes type II, and to increase lean
CC mass and insulin sensitivity. They may be used in conjunction with
CC insulin, hypoglycaemics, antihypertensives, diuretics, appetite
CC suppressants etc. These suspensions are stable and active at
CC physiological pH and are ready-for-use formulations that do not require
CC freezing or freeze drying. As they are very concentrated, only small
CC volumes are required and they provide a sustained-release effect, with
CC increased potency and reduced frequency of injection.
XX Sequence 374 AA;

Query Match 93.3%; Score 1245; DB 19; Length 374;
Best Local Similarity 74.4%; Pred. No. 4.5e-35;
Matches 233; Conservative 4; Mismatches 7; Indels 69; Gaps 2;
QY 1 MDKTHCTCPAPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKENWYV 60
DB 1 MDKTHCTCPAPAPPELLGGPSVFLPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKENWYV 60
QY 61 DGVEVHNAKTPREEQNSTYRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKAK 120
DB 61 DGVEVHNAKTPREEQNSTYRVVSVLTVLHODWLNKGEYKCKVSNKALPAPIEKTISKAK 120
QY 121 KGQREPOVYTLPPSRDELTKNOVSLCLVKGFPSPDIWESNGQPNENYKTTTPPVLD 180
DB 121 KGQREPOVYTLPPSRDELTKNOVSLCLVKGFPSPDIWESNGQPNENYKTTTPPVLD 180
QY 181 SDGSFFLYSKLTVDKSRWQOGNVSFSCVMHEALHNHYTQKSLSLSPGK----- 228
DB 181 SDGSFFLYSKLTVDKSRWQOGNVSFSCVMHEALHNHYTQKSLSLSPGKVPQKVQDDTKT 240
QY 229 -----DWLKAFFYDKV----- 238
DB 241 LKTIIVTRINDISHTQSVSSKQKVTGLDIPGLHPILTLSKMDQTLAVYQQIILTSWPSRN 300
QY 239 -----AEKLKE 244
DB 301 VIQISNDLENLRD 313

RESULT 11
AAW49075
ID AAW49075 standard; Protein; 374 AA.
XX AAW49075;
AC AAW49075;
XX 18-NOV-1998 (first entry)
DE Recombinant human MetFc-OB variant 2 protein.
XX
KW Recombinant human MetFc-OB variant 2 protein; chimeric; immunoglobulin;
KW high blood lipid level; arterial sclerosis; stroke; Fc-OB fusion protein;
KW diabetes.
XX


```

QY 1 MDKTHCTPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60
DB 1 MDKTHCTPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60
QY 61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 120
DB 61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 120
QY 121 KGPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVFNESNGQPNENNYKTTTPPVLD 180
DB 121 KGPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVFNESNGQPNENNYKTTTPPVLD 180
QY 181 SDGSFFLYSKLTVDKSRWQQGNVFNCSVMHEALHNHYTQKSLSLSPGK----- 228
DB 181 SDGSFFLYSKLTVDKSRWQQGNVFNCSVMHEALHNHYTQKSLSLSPGKGGGIEGPTLR 240
QY 229 DWLKA 233
DB 241 QWLAA 245

RESULT 8
ID AAM52156 standard; Protein; 731 AA.
AC AAM52156;
XX AC
DT 05-FEB-2002 (first entry)
XX DT
DE Humanised HMFG-1 heavy chain/DNase I fusion protein 1.
XX DE
KW Humanised monoclonal antibody; polymorphic epithelial mucin; PEM1;
KW cytotoxic; endonuclease; DNase I; human; cytostatic; cancer; apoptosis.
XX KW
OS Homo sapiens.
OS Synthetic.
XX OS
PN WO200174905-A1.
XX PN
PD 11-OCT-2001.
XX PD
PF 26-MAR-2001; 2001WO-GB01324.
XX PF
PR 03-APR-2000; 2000GB-0008049.
XX PR
PR 02-OCT-2000; 2000US-237159P.
XX PR
XX (ANTI-) ANTISOMA RES LTD.
PA PA
PI Young RJ;
XX PI
DR WPI; 2001-662969/76.
XX DR
PT Novel compound used to treat cancer has target cell-specific portion
PT comprising humanised monoclonal antibody having specificity for
PT polymorphic epithelial mucin, and cytotoxic portion having
PT endonucleolytic activity
XX PT
PS Claim 20; Figure 7; 176pp; English.
XX PS
CC The invention relates to a compound which comprises a target
CC cell-specific portion, comprising an humanised monoclonal antibody,
CC having specificity for polymorphic epithelial mucin (PEM) or its antigen
CC binding fragment and a cytotoxic portion having endonucleolytic activity,
CC exemplified by AAM52154-AAM52168 and encoded by ABA02682-ABA02728. The
CC compound has cytostatic activity useful for treating cancer and acting as
CC a potential inducer of apoptosis.
XX CC
SQ Sequence 731 AA;
XX SQ
Query Match 93.4%; Score 1246; DB 22; Length 731;
Best Local Similarity 48.6%; Pred. No. 1.9e-34;
Matches 233; Conservative 6; Mismatches 5; Indels 235; Gaps 3;

```

```

QY 2 DKTHCTPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 61
DB 241 DKTHCTPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 300
QY 62 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 121
DB 301 GVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 360
QY 122 GPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVFNESNGQPNENNYKTTTPPVLD 181
DB 361 GPREPQVYTLPPSRDELTKNOVSLTCLVKGFPSPDIAVFNESNGQPNENNYKTTTPPVLD 420
QY 182 DGSFFLYSKLTVDKSRWQQGNVFNCSVMHEALHNHYTQKSLSLSPGK----- 228
DB 421 DGSFFLYSKLTVDKSRWQQGNVFNCSVMHEALHNHYTQKSLSLSPGKGGGLKIAAFNIQ 480
QY 229 ----- 228
DB 481 TFGETKMSNATLVSYIVQILSRYDIALVOEVRDLSHTAVGKLLDNLNQDAPDTYHYVNSE 540
QY 229 ----- 228
DB 541 PLGRNSYKERYLFVYRPDQVSAVDYDDGCEPCGNDTFNREPAIVRFSETEVREFA 600
QY 229 ----- 228
DB 601 IVPLHAAPGDAVAEIDALYDVYLDVQEKWGLDVMLMGDFNAGCSYVRPSQWSSIRLWTS 660
QY 229 ---DWL-----KAFYDKV-----AEKLUKEA 245
DB 661 PTFQWLIPDSADTTATPTTHCAYDRIVVAGMLLRGAVVPDSALPFNFQAAAYGLSDQLAAQ 719

RESULT 9
AAM52159
ID AAM52159 standard; Protein; 741 AA.
XX AC
XX AAM52159;
XX AC
DT 05-FEB-2002 (first entry)
XX DT
DE Humanised HMFG-1 heavy chain/DNase I fusion protein 4.
XX DE
KW Humanised monoclonal antibody; polymorphic epithelial mucin; PEM1;
KW cytotoxic; endonuclease; DNase I; human; cytostatic; cancer; apoptosis.
XX KW
OS Homo sapiens.
OS Synthetic.
XX OS
PN WO200174905-A1.
XX PN
PD 11-OCT-2001.
XX PD
PF 26-MAR-2001; 2001WO-GB01324.
XX PF
PR 03-APR-2000; 2000GB-0008049.
XX PR
PR 02-OCT-2000; 2000US-237159P.
XX PR
XX (ANTI-) ANTISOMA RES LTD.
PA PA
PI Young RJ;
XX PI
DR WPI; 2001-662969/76.
XX DR
PT Novel compound used to treat cancer has target cell-specific portion
PT comprising humanised monoclonal antibody having specificity for
PT polymorphic epithelial mucin, and cytotoxic portion having
PT endonucleolytic activity
XX PT
PS Claim 20; Figure 10; 176pp; English.
XX PS
CC The invention relates to a compound which comprises a target

```


CC neuroprotective activities. (I) can be used as a therapeutic or
CC prophylactic agent as well as for screening purposes. (I) is useful for
CC diagnosing diseases characterised by dysfunction of their associated
CC protein of interest, for identifying normal or abnormal proteins of
CC interest, as a part of diagnostic kit to detect the presence of their
CC proteins of interest in a biological sample. Additionally, (I) is useful
CC for treating inflammatory and autoimmune diseases, tumour growth, cancer,
CC rheumatoid arthritis, diabetic retinopathy, obesity, sleep disorders,
CC infertility, and neurological degenerative diseases. (I), comprising
CC EPO-mimetic compounds are useful for treating disorders characterised by
CC low red blood cell levels such as anaemia. The TPO-mimetic comprising
CC compounds are useful for treating conditions that involve an existing
CC megakaryocyte/platelet deficiency or an expected megakaryocyte/platelet
CC deficiency, such as thrombocytopenia, aplastic anaemia, metastatic
CC tumour which result in thrombocytopenia, systemic lupus erythematosus,
CC and Fanconi's syndrome. ABB72403 to ABB73426 and ABL35695 to ABL35777
CC represent amino acid and nucleic acid sequences used in the
CC exemplification of the present invention.

XX Sequence 248 AA;

Query Match 93.4%; Score 1246; DB 23; Length 248;
Best Local Similarity 93.9%; Pred. No. 1.6e-35;
Matches 230; Conservative 3; Mismatches 6; Indels 6; Gaps 1;

QY 1 MDKTHTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60
DB 1 MDKTHTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60
QY 61 DGVEVHNATKPREQYNSTYRVSVLTVLHODWLNKGEYCKVKSNKALPAPIEKTISK 120
DB 61 DGVEVHNATKPREQYNSTYRVSVLTVLHODWLNKGEYCKVKSNKALPAPIEKTISK 120
QY 121 KGQPREQVYTLPPSRDELTKNOVSLTCLVKGYFSPSDIAVEWESNGQPENNYKTTTPVLD 180
DB 121 KGQPREQVYTLPPSRDELTKNOVSLTCLVKGYFSPSDIAVEWESNGQPENNYKTTTPVLD 180
QY 181 SGGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKQSLSPGK-----DWLKAF 234
DB 181 SGGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKQSLSPGKGGGGFETWPGY 240
QY 235 YDKVA 239
DB 241 WQPYA 245

RESULT 5
AAB16959
ID AAB16959 standard; Protein; 268 AA.
XX AAB16959;
XX AAB16959;
XX 31-OCT-2000 (first entry)
XX FC-TMP-TMP protein sequence SEQ ID NO:8.

XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW vascular endothelial growth factor; matrix metalloproteinase;
KW asthma; thrombosis; pharmaceutical.

XX Homo sapiens.
OS Synthetic.
XX WO200024782-A2.
XX 04-MAY-2000.
XX 25-OCT-1999; 99WO-US25044.

PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX (AMGE-) AMGEN INC.
XX Feige U, Liu C, Cheetham J, Boone TC;
XX WPI; 2000-350702/30.
XX DR N-PSDB; AAB69445.
XX Novel composition of matter comprising an Fc domain and
XX pharmacologically active peptides, useful for treating cancer and
XX autoimmune diseases -
XX Example 2; Page 182-183; 608pp; English.

XX The present invention describes composition of matter (I) comprising an
XX Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
XX (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
XX independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
XX -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
XX where P1, P2, P3, and P4 = are each independently sequences of
XX pharmacologically active peptides; L1, L2, L3, and L4 = are each
XX independently linkers; and a, b, c, d, e, and f = are each independently
XX 0 or 1, provided that at least 1 of a and b is 1. The composition can
XX have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
XX activities. DNAs, vectors and host cells from the present invention can
XX be used for producing pharmaceutical compositions. The compositions are
XX useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
XX The use of an Fc domain (rather than a Fab domain) can provide a longer
XX half-life or incorporate functions such as Fc receptor binding, protein
XX A binding, complement fixation, and possibly placental transfer. AAB69443
XX to AAB69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
XX sequences used in the exemplification of the present invention.

XX Sequence 268 AA;

Query Match 93.4%; Score 1246; DB 21; Length 268;
Best Local Similarity 94.3%; Pred. No. 1.9e-35;
Matches 231; Conservative 0; Mismatches 2; Indels 12; Gaps 1;

QY 1 MDKTHTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60
DB 1 MDKTHTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVDVSHEDPEVKFNWYV 60
QY 61 DGVEVHNATKPREQYNSTYRVSVLTVLHODWLNKGEYCKVKSNKALPAPIEKTISK 120
DB 61 DGVEVHNATKPREQYNSTYRVSVLTVLHODWLNKGEYCKVKSNKALPAPIEKTISK 120
QY 121 KGQPREQVYTLPPSRDELTKNOVSLTCLVKGYFSPSDIAVEWESNGQPENNYKTTTPVLD 180
DB 121 KGQPREQVYTLPPSRDELTKNOVSLTCLVKGYFSPSDIAVEWESNGQPENNYKTTTPVLD 180
QY 181 SGGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKQSLSPGK-----228
DB 181 SGGSFFLYSKLTVDKSRWQGNVFCVSMHEALHNHYTKQSLSPGKGGGGIEGPTLR 240
QY 229 DWLKA 233
DB 241 QWLAA 245

RESULT 6
ABB73412
ID ABB73412 standard; Protein; 268 AA.
XX ABB73412;
XX ABB73412;
XX 05-APR-2002 (first entry)
XX FC-TMP-TMP amino acid SEQ ID NO:8.
XX Modified peptide; mimetic; Fc domain; fusion; Immunoglobulin G; IgG;
KW

Db 181 SDGSFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPOKGGGGGEGPTLR 240
 QY 229 DWLKA 233
 II I
 Db 241 QWLAA 245

RESULT 3
 ID AAB17953
 XX AAB17953 standard; Protein; 248 AA.
 AC AAB17953;
 XX 31-OCT-2000 (first entry)
 XX
 DE FC-IL-1 antagonist fusion protein sequence SEQ ID NO:1060.
 XX
 KW Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
 KW autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
 KW immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
 KW MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
 KW cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
 KW vascular endothelial growth factor; matrix metalloproteinase;
 KW asthma; thrombosis; pharmaceutical.
 XX
 OS Synthetic.
 XX
 PN WO200024782-A2.
 XX
 PD 04-MAY-2000.
 XX
 PF 25-OCT-1999; 99WO-US25044.
 XX
 PR 23-OCT-1998; 98US-0105371.
 PR 22-OCT-1999; 99US-042082.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Feige U, Liu C, Cheetham J, Boone TC;
 XX
 DR WPI; 2000-350702/30.
 DR N-PSDB; AAA69503.
 XX
 PT Novel composition of matter comprising an Fc domain and
 PT pharmacologically active peptides, useful for treating cancer and
 PT autoimmune diseases -
 XX
 PS Example 5; Page 574-575; 608pp; English.
 XX
 CC The present invention describes composition of matter (I) comprising an
 CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
 CC (X1)-a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
 CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
 CC -(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
 CC where P1, P2, P3, and P4 = are each independently sequences of
 CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
 CC independently linkers; and a, b, c, d, e, and f = are each independently
 CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
 CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
 CC activities. DNAs, vectors and host cells from the present invention can
 CC be used for producing pharmaceutical compositions. The compositions are
 CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
 CC The use of an Fc domain (rather than a Fab domain) can provide a longer
 CC half-life or incorporate functions such as Fc receptor binding, protein
 CC A binding, complement fixation, and possibly placental transfer. AAA69443
 CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
 CC sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 248 AA;

Query Match 93.4%; Score 1246; DB 21; Length 248;
 Best Local Similarity 93.9%; Pred. No. 1.6e-35;
 Matches 230; Conservative 3; Mismatches 6; Indels 6; Gaps 1;

QY 1 MDKTHCTPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTVVVDVSHEDPEVKFNMY 60
 II I
 Db 1 MDKTHCTPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTVVVDVSHEDPEVKFNMY 60
 QY 61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
 II I
 Db 61 DGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKA 120
 QY 121 KGQPREPOVYTLPPSRDELTKNQVSLTCLIVKGFPYPSDIAVWESNGQPENNYKTTPVLD 180
 II I
 Db 121 KGQPREPOVYTLPPSRDELTKNQVSLTCLIVKGFPYPSDIAVWESNGQPENNYKTTPVLD 180
 QY 181 SDGSFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPOKGGGGGEGPTLR 240
 II I
 Db 181 SDGSFLYSLKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSPOKGGGGGEGPTLR 240
 QY 235 YDKVA 239
 II I
 Db 241 WQPYA 245

RESULT 4
 ID ABB73421
 XX ABB73421 standard; Protein; 248 AA.
 AC ABB73421;
 XX
 DT 05-APR-2002 (first entry)
 XX
 DE Fc-interleukin 1 (IL-1) antagonist fusion nucleic acid SEQ ID NO:1059.
 XX
 KW Modified peptide; mimetic; Fc domain; fusion; immunoglobulin G; IgG;
 KW EPO; erythropoietin; TPO; tumour necrosis factor alpha inhibitor;
 KW TNF-alpha inhibitor; interleukin 1 antagonist; IL-1 antagonist; TMP;
 KW TPO mimetic peptide; EPO mimetic peptide; EMP; VEGF antagonist;
 KW MMP inhibitor; antiinflammatory; antitumour; immunosuppressive;
 KW cytostatic; antirheumatic; antiarthritic; antidiabetic; ophthalmological;
 KW antianaemic; anorectic; antinfertility; haemostatic; dermatological;
 KW neuroprotective; inflammatory disease; autoimmune disease; tumour growth;
 KW cancer; rheumatoid arthritis; diabetic retinopathy; infertility; obesity;
 KW sleep disorder; neurological degenerative disease; anaemia;
 KW thrombocytopaenia; metastatic tumour; systemic lupus erythematosus;
 KW Fanconi's syndrome.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO200183525-A2.
 XX
 PD 08-NOV-2001.
 XX
 PF 02-MAY-2001; 2001WO-US14310.
 XX
 PR 03-MAY-2000; 2000US-0563286.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Feige U, Liu C, Cheetham JC, Boone TC, Gudas JM;
 XX
 DR WPI; 2002-130313/17.
 DR N-PSDB; ABL35771.
 XX
 PT Novel vehicle-peptide molecule or its multimers useful for treating
 PT inflammatory and autoimmune diseases, cancer, rheumatoid arthritis,
 PT diabetic retinopathy, obesity, sleep disorders and infertility -
 XX
 PS Example 5; Fig 21A-B; 176pp; English.
 XX
 CC The present invention describes a vehicle-peptide molecule (I) or its
 CC multimers. (I) can have antiinflammatory, antitumour, immunosuppressive,
 CC cytostatic, antirheumatic, antiarthritic, antidiabetic, ophthalmological,
 CC antianaemic, anorectic, antinfertility, haemostatic, dermatological and